



THESE

**présentée pour obtenir le titre de Docteur
de l'Ecole Nationale Supérieure
des Télécommunications**

Spécialité : Electronique et Communications

Hélio MAGALHAES DE OLIVEIRA

**SUR LA THEORIE DES SYSTEMES
DE MODULATION CODÉE POUR
LA TRANSMISSION NUMERIQUE À DEBIT ELEVÉ**

TELECOM Paris 92 E 001

SOUTENUE LE 14 FEVRIER 1992 DEVANT LE JURY COMPOSÉ DE

Michel JOINDOT

Président et Rapporteur

Gérard BATTAIL

Examineurs

Marc ROUANNE

Hikmet SARI

Jean-Claude BIC

Rapporteur

Sur la Théorie des Systèmes de Modulation Codée pour la Transmission Numérique à Débit Elevé.

Le début est consacré à un aperçu historique et le reste à des contributions personnelles. Le problème de reconstituer la phase de la porteuse a été envisagé pour les Modulations Codées en Treillis (MCT), aboutissant à une MCT différentiellement cohérente. Un étiquetage binaire optimal des symboles d'une constellation carrée a été introduit et une famille de constellations en croix généralisées étudiée. Des bornes fines de leurs performances ont été établies et leur efficacité spectrale a été évaluée.

Utilisant le théorème de géométrie des nombres de Minkowski-Hlawka, on a démontré un théorème de capacité pour le canal à bruit additif, blanc et gaussien, établissant ainsi l'existence de moyens déterministes pour approcher la capacité.

Des bornes sur les performances des codes de réseau ont été obtenues. Une amélioration de l'approximation continue de Forney pour estimer la puissance moyenne normalisée à 2 dimensions a été suggérée. On a analysé le comportement de ces bornes pour un grand nombre de dimensions et conjecturé que s'approcher de la capacité est lié à la recherche de bons recouvrements plutôt qu'à celle d'empilements denses.

On a appliqué le raisonnement de Shannon basé sur le codage aléatoire à des constellations de nombre de dimensions fini. Les spectres de distance euclidienne de telles constellations ont été calculés. Un phénomène de durcissement de la distance a été trouvé. On a montré que l'affirmation "tous les codes sont bons" reste vraie pour les constellations codées.

ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS

CENTRE DOCUMENTAIRE - IMPRIMERIE - 46 rue Barrault - 75634 Paris Cedex 13

Téléphone (1) 45 81 77 77 - Télécopie (1) 45 89 79 06 - Telex SUPTLCM 200160F

© 1992 FRANCE TELECOM - RCS PARIS B 380 129 866



DIRECTION DE L'ENSEIGNEMENT SUPERIEUR

THESE

SPECIALITÉ : ELECTRONIQUE ET COMMUNICATIONS

HÁLIO MAGALHÃES DE OLIVEIRA

TELECOM Paris 92 E 001



THESE

**présentée pour obtenir le titre de Docteur
de l'Ecole Nationale Supérieure
des Télécommunications**

Spécialité : Electronique et Communications

Hélio MAGALHAES DE OLIVEIRA

**SUR LA THEORIE DES SYSTEMES
DE MODULATION CODÉE POUR
LA TRANSMISSION NUMERIQUE À DEBIT ELEVÉ**

TELECOM Paris 92 E 001

SOUTENUE LE 14 FEVRIER 1992 DEVANT LE JURY COMPOSÉ DE

Michel JOINDOT

Président et Rapporteur

Gérard BATAIL

Examineurs

Marc ROUANNE

Hikmet SARI

Jean-Claude BIC

Rapporteur



THESE

**présentée pour obtenir le titre de Docteur
de l'Ecole Nationale Supérieure
des Télécommunications**

Spécialité : Electronique et Communications

Hélio MAGALHAES DE OLIVEIRA

**SUR LA THEORIE DES SYSTEMES
DE MODULATION CODÉE POUR
LA TRANSMISSION NUMERIQUE À DEBIT ELEVÉ**

TELECOM Paris 92 E 001

SOUTENUE LE 14 FEVRIER 1992 DEVANT LE JURY COMPOSÉ DE

Michel JOINDOT

Président

Gérard BATTAIL

Examineurs

Marc ROUANNE

Hikmet SARI

Jean-Claude BIC

Rapporteurs

Michel JOINDOT

A mes parents, sans l'affection et le soutien desquels je n'aurais certainement jamais accompli cette tâche. Dans les pages qui suivent, qu'ils trouvent une concrétisation supplémentaire de leurs effort. Qu'ils soient em même temps assurés de mon infinie tendresse.

REMERCIEMENTS

En tout premier lieu, qu'il me soit permis d'exprimer ma haute et respectueuse considération à M. le Professeur M. JOINDOT de l'Université de Rennes I, qui a bien voulu me faire l'honneur de présider le jury de soutenance.

Je voudrais aussi exprimer tous mes remerciements à Messieurs H. SARI (chef du département d'Etudes et Développement SAT, professeur associé) et M. ROUANNE (chef de group MATRA COM), d'avoir bien voulu prendre part à ce jury et ainsi que de l'intérêt qu'ils ont porté à mon travail.

J'exprime également ma respectueuse reconnaissance envers MM. J.C. BIC (chef du départ. RPE/ETP du CNET Paris B) et M. JOINDOT, qui ont accepté de participer au jury en tant que rapporteurs et dont les remarques m'ont été précieuses.

Je témoigne ma reconnaissance à M. le Directeur Scientifique de l'ENST, M C. GUEGUEN, pour m'avoir offert la possibilité de réaliser ce mémoire. Qu'il soit assuré de mes respectueux sentiments.

Je suis, d'autre part, très redevable envers MM. P. CHEMOIL (CNET - ATR) et J. BERNOUSSU (LAAS Toulouse) qui m'ont fait découvrir TELECOM PARIS et qui m'ont gentiment aidé dans les démarches d'inscription.

Je tiens à exprimer ici tous mes remerciements et ma profonde reconnaissance à M. le Professeur G. BATTAIL, pour la confiance qu'il m'a témoignée en acceptant de diriger ce travail et pour l'intérêt *permanent* qu'il a manifesté à l'égard de cette recherche. Chercheur au sens noble du mot, le Prof. Battail a bien su me faire profiter de sa longue expérience et de ses idées. Ses remarques, avec beaucoup d'esprit critique et de rigueur scientifique, ont contribué à de *nombreux* points détaillés ici. Ses critiques sur la validité du critère de distance minimale pour les codes en bloc dont la longueur devient arbitrairement grande ont eu une profonde influence sur ma pensée dans le domaine. J'ai particulièrement apprécié qu'il ait su à la fois guider mon travail avec une compétence jamais démentie, et me laisser beaucoup de liberté dans le choix de mes orientations. En outre, je le remercie également pour les encouragements qu'il a su me prodiguer pendant toute la durée de cette étude, et pour son aide fort précieuse lors de la rédaction de publications et du manuscrit.

Qu'il trouve ici l'expression de ma sincère reconnaissance et de mon respect.

Je ne saurais manquer de remercier Mr. G. KAWAS KALEH, Maître de Conférences, pour l'attention toute particulière qu'il a portée à l'élaboration de cette thèse. Ses conseils précieux m'ont, en quelques discussions sur la modulation codée en treillis, beaucoup apporté et certainement ont contribué à l'aboutissement de mon travail. Je remercie aussi Mr. J.C. BELFIORE qui m'a fait l'honneur de suivre avec attention le déroulement de cette étude et dont le sympathique accueil a rendu mon séjour à l'ENST très agréable.

Je souhaiterais exprimer aussi ma gratitude au Prof. V.C. DA ROCHA pour m'avoir entraîné dans la Recherche. Je lui sais gré de m'avoir initié à la théorie du codage. C'est aussi avec plaisir que je remercie le Prof. F.M. CAMPELLO DE SOUZA qui m'a fait partager son enthousiasme pour la recherche et qui m'a poussé à faire un doctorat en France. En témoignage de ma sincère amitié.

Le travail de recherche présenté dans ce mémoire a pu être réalisé grâce au financement de l'Université Fédérale de Pernambouc (UFPE-Brésil). Par ailleurs, il a aussi bénéficié de l'aide financière supplémentaire de l'ARECOM. Je leur adresse mes très sincères remerciements.

J'ai beaucoup utilisé, parfois avec excès, les innombrables services de la documentation technique de l'ENST. J'adresse un grand merci à l'intention de ses membres, en particulier à Mme A. VIDAL, pour leur efficacité et leur compréhension.

Je tiens à exprimer ma gratitude à tout le personnel du Dép. COM, en particulier à Mme J. BATS, pour leur gentillesse et leur disponibilité.

J'adresse des vifs remerciements à mon compagnon, M. ZHANG weidong (張卫东), avec qui j'ai spécialement conjugué les efforts pendant mes recherches. Je dois beaucoup aux collègues du Dép. COM, avec lesquels j'ai parfois été amené à collaborer, particulièrement à M. Anders PETERSSON et M. YUAN Chen (袁晨), sans oublier mon cher camarade Théophanis KALLIACOUDAS (Καλλιακουδας θεοφανης). Je conserve aussi un excellent souvenir de ma collaboration avec Mr. Byong-Ki MIN (Dép. ELE) avec qui j'ai eu de fructueuses discussions.

Qu'ils en soient tous vivement remerciés.

Je suis très reconnaissant à tous mes collègues et amis de l'ENST qui m'ont rendu agréable la poursuite de mes travaux et dont l'amical soutien m'a profondément touché. Parmi mes camarades, il m'est impossible de citer tous ceux qui y ont contribué: Je salue notamment Mme Kitty LEUWEEN-BOULLE et Mme Patricia MENDEZ, aussi bien que M. Olivier PEZE (PZ). J'en sais gré à tous. Je témoigne aussi mon amitié à Francisco et Maryse PEREIRA. Qu'ils se sachent assurés de mon affection. Leur contribution indirecte à cette thèse est loin d'être négligeable. Je veux tout aussi chaleureusement remercier Mme Icilia de OLIVEIRA pour les encouragements qu'elle m'a prodigués. En témoignage de mon attachement.

Je souhaite enfin exprimer combien ce séjour en France et au sein du Dép. COM m'a été inoubliable. Il m'a permis d'accéder à bien des enrichissements, tant d'un point de vue professionnel que personnel. Que tous soient assurés de ma gratitude et de mon amitié.

Je ne saurais clore sans adresser le plus grand Merci à ma famille: Merci tendrement à mon épouse Néréide, et à mes enfants Raquel et Raphaël, d'être à mes côtés.

RESUME

La première partie de cette thèse est consacrée à un aperçu historique préalable et la seconde partie à des contributions personnelles. Les problèmes de récupération de la phase de la porteuse en modulation codée en treillis (MCT) ont fait l'objet d'une étude aboutissant à une MCT différentiellement cohérente. On a introduit un étiquetage binaire optimal des symboles des constellations carrées et étudié une famille de constellations en croix généralisées. Des bornes fines de leurs performances ont été développées et leur efficacité spectrale a été évaluée.

En s'appuyant sur le théorème de géométrie des nombres de Minkowski-Hlawka, on a démontré un théorème de capacité pour les codes de réseau sur le canal à bruit additif, gaussien et blanc, y compris son assertion négative. On a établi ainsi l'existence de moyens déterministes (codes fortement structurés) pour s'approcher de la capacité.

Des bornes sur les performances des codes de réseau ont été établies. On a suggéré une amélioration de l'approximation continue de Forney utilisée pour estimer la puissance moyenne normalisée à deux dimensions. On a analysé en particulier le comportement de ces bornes pour un nombre de dimensions élevé, ce qui a conduit à la conjecture que le problème de s'approcher de la capacité est lié davantage à la recherche de bons recouvrements qu'à celle d'empilements denses.

L'accent a ensuite été mis sur le codage-modulation de constellations à nombre de dimensions fini. Dans cette perspective, on a appliqué à la modulation codée le raisonnement de Shannon basé sur le codage aléatoire et la critique du critère de distance minimale revendiquée par Battail. Les spectres de distance euclidienne de divers codages-modulations aléatoires ont été évalués. On a mis en évidence un phénomène de "durcissement" de la distance normalisée dans les espaces euclidiens à grand nombre de dimensions, ce qui équivaut à une borne de Varshamov-Gilbert sur la distance euclidienne. En outre, on a démontré que l'énoncé "tous les codes sont bons" est vérifié aussi pour les constellations codées.

Les investigations menées dans cette étude aboutissent donc à proposer un meilleur aperçu du comportement asymptotique de la modulation codée.

ABSTRACT

The first part of this thesis is devoted to a preliminary historical overview while the second one contains personal contributions. The problem of carrier phase recovery is addressed in the case of trellis-coded modulation (TCM), resulting in a differentially coherent TCM. An optimum binary labelling of the symbols in square constellations has been introduced and a family of generalized cross constellations has been studied. Tight bounds on their performance were developed and their spectral efficiency was evaluated.

By means of the Minkowski-Hlawka theorem of the geometry of numbers, we proved a capacity theorem for lattice codes over the additive white Gaussian noise channel, including its negative statement (strong converse). Therefore the existence of deterministic means (i.e., highly structured codes) to approach capacity was established.

Bounds on the performance of lattice codes were obtained. An improvement of Forney's continuous approximation used in order to estimate the mean normalized two-dimensional power was suggested. We analyzed in particular the behaviour of these bounds for high dimensionalities, which led us to conjecture that the problem of approaching capacity is connected with searching for good coverings rather than for dense packings.

Emphasis was then put on coding-modulation for a finite number of dimensions. In this context, we applied Shannon's random coding argument as well as Battail's criticism of the minimum distance criterion. Euclidean distance spectra of several coding-modulation schemes were evaluated. A phenomenon of normalized distance "hardening" was found when the dimensionality increases, which can be interpreted as a kind of Varshamov-Gilbert bound on the Euclidean distance. Moreover, we showed that the statement "all codes are good" still holds for the coded constellations.

The investigations carried out in this study thus result in proposing a better overview on the asymptotic behaviour of coded modulation.

AVANT PROPOS

J'ai eu le privilège de conduire cette recherche au sein d'une des plus prestigieuses écoles de Télécommunications d'Europe et de travailler avec certains de ses plus renommés chercheurs. Cette thèse est consacré à l'étude théorique des systèmes de communications numériques à haut débit destinés à des canaux appartenant à la région limitée en bande passante, ce qui constitue la grande majorité des voies réelles. L'actualité, voire l'importance du sujet paraît incontestable: il s'agit de la plus grande révolution de la théorie (et de la pratique) des télécommunications depuis la publication des travaux fondamentaux de Shannon en 1948.

La difficulté que j'ai éprouvée à la rédiger découle partiellement de la liberté d'action dont j'ai bénéficié pour la mener. J'ai même eu du mal à choisir un titre approprié: celui que j'ai finalement adopté n'élucide pas complètement le contenu du mémoire. Cependant, sa diversité reflète ma pensée. Mon effort a été concentré sur la création, rechercher et créer étant à peu près synonymes. Les résultat y sont présentés avec rigueur, *ma non troppo*. Ma réponse aux ingénieurs-mathématiciens "trop formels" est celle du génial O. Heaviside: "Eh bien, qu'en dites-vous? Doit-on refuser à se nourrir parce qu'on ne comprend pas tout le processus digestif?" (il n'y aura point de réponse à ceux qui m'accusent d'un formalisme excessif).

Parmi quelques contributions faites dans le (beau) domaine de la modulation codée, deux points sont particulièrement mis en relief, à savoir: tout d'abord la démonstration de l'existence de bons codes de réseau (*lattice codes*) basée sur l'application du théorème de Minkowski de la géométrie des nombres. Ensuite, l'application à la modulation codée du concept de codage aléatoire, base du raisonnement qui conduit à la démonstration du théorème fondamental du codage de canal.

Les investigations effectuées dans le cadre de cette thèse sont destinées à fournir un meilleur aperçu du mécanisme du codage de jeux de signaux multidimensionnels et des performances qui en résultent. Les conclusions rejoignent clairement les idées du Prof. Battail à propos de sa remise en question de la validité du critère de distance minimale et de l'existence de moyens déterministes de codage imitant le codage aléatoire.

Quoi qu'il en soit, j'ose dire que l'originalité y est présente. *Alea ajact est*.

H. Magalhães (Magellan) de Oliveira.

TABLE DES MATIERES

| | |
|---|----|
| INTRODUCTION GENERALE | 1 |
| <u>TRAVAUX ANTERIEURS</u> | 9 |
| CHAPITRE I | |
| INTRODUCTION AUX SYSTEMES DE MODULATION CODEE ..11 | |
| I.1 <u>BREFS RAPPELS SUR L'HISTOIRE DES COMMUNICATIONS</u> ...13 | |
| I.1.1 ANCIENNE HISTOIRE DES TRANSMISSIONS NUMERIQUES | 13 |
| I.1.2 L'IDEE ET L'HISTOIRE DE LA MODULATION CODEE | 16 |
| I.1.3 L'EVOLUTION DES MODEMS | 19 |
| I.1.4 DE LA POSSIBILITE DE TROUVER DE BONS CODES LONGS | 27 |
| I.2. <u>APERÇU SUR LES CODES D'UNGERBOECK</u> | 30 |
| I.3. <u>ETAT DE L'ART</u> | 37 |
| I.3.1 GAIN POTENTIEL DE LA MODULATION CODEE..... | 37 |
| I.3.2 MODULATION CODEE EN TREILLIS (MCT) | 42 |
| I.3.3 MODULATION CODEE EN BLOC (MCB) | 48 |
| I.4. <u>THEORIE MODERNE DE LA MODULATION CODEE:</u> <u>QUELQUES SUJETS AVANCES</u> | 52 |
| I.4.1 MODULATION CODEE MULTIDIMENSIONNELLE | 52 |
| I.4.2 CODES DE RESEAU ET DE CLASSES LATERALES | |
| I.4.2.1 RESEAUX | 54 |
| I.4.2.2 CODES DE CLASSE LATERALES | 66 |

TRAVAUX PERSONNELS77

CHAPITRE II

CERTAINS ASPECTS DE LA MODULATION ET DU CODAGE..79

II.1. CONSTELLATIONS EN CROIX GENERALISEES ET CANAL SECONDAIRE OPPORTUNISTE81

II.1.1 INTRODUCTION82

II.1.2 FAST DETECTION ALGORITHM FOR SQUARE QAM CONSTELLATIONS83

II.1.3 PERFORMANCE OF GENERALIZED CROSS CONSTELLATIONS86

II.1.4 OPPORTUNISTIC SECONDARY CHANNEL96

II.1.5 CONCLUDING REMARKS99

II.2. PROBLEMES LIES A LA RECUPERATION DE PORTEUSE (MCT): MODULATION CODEE DIFFERENTIELLEMENT COHERENTE107

II.2.1 INTRODUCTION109

II.2.2 DESIGN OF TRELLIS-CODE M-ary DPSK MODULATION111

II.2.3 TRELLIS-CODED DIFFERENTIAL MODULATION119

II.2.4 CONCLUSIONS124

CHAPITRE III

A PROPOS DE L'EXISTENCE DE BONS CODES DE RESEAU. 125

III.1. A CAPACITY THEOREM FOR LATTICE CODES.....127

III.1.1 INTRODUCTION128

III.1.2 BOUNDS ON THE ERROR PROBABILITY129

III.1.3 SPHERE HARDENING THEOREMS132

III.1.4 CODING THEOREMS FOR LATTICE CODES134

III.1.5 PERFORMANCE OF LATTICE CODES OVER THE AWGN CHANNEL137

III.1.6 CONCLUSIONS138

| | |
|---|------------|
| III.2. PERFORMANCE OF LATTICE CODES OVER THE GAUSSIAN CHANNEL | |
| III.2.1 INTRODUCTION | 141 |
| III.2.2 BOUNDS ON THE N-DIMENSIONAL ERROR PROBABILITY | 142 |
| III.2.3 LATTICE POINTS CONFINED IN BOUNDED REGIONS | 146 |
| III.2.4 GAINS, VORONOI REGIONS AND MORE ON THE BEHAVIOUR OF LATTICES | 150 |
| III.2.5 BY WAY OF INTERPRETATION | 153 |
| III.2.6 CONCLUSIONS AND SUMMARY | 155 |
| CHAPITRE IV | |
| THEORIE DE LA MODULATION CODEE ALEATOIRE. | 163 |
| IV.1. PERFORMANCE DU CODAGE-MODULATION POUR LE CANAL GAUSSIEN: | 165 |
| IV.1.1 INTRODUCTION | 166 |
| IV.1.2 A POWERFUL TOOL: THE RANDOM CODING | 167 |
| IV.1.3 PERFORMANCE PREDICTION USING RANDOM CODING | 168 |
| IV.1.4 PRESENCE OF RAYLEIGH FLUCTUATIONS | 170 |
| IV.1.5 CONCLUSION | 173 |
| IV.2. THE RANDOM CODED MODULATION: PERFORMANCE AND EUCLIDEAN DISTANCE SPECTRUM EVALUATION..... | 175 |
| IV.2.1 INTRODUCTION | 176 |
| IV.2.2 BOUNDS ON THE ERROR PROBABILITY AND PARTIAL CUT-OFF RATE | 177 |
| IV.2.3 EUCLIDEAN DISTANCE SPECTRUM OF RANDOM CODES ... | 186 |
| IV.2.4 MORE ON THE ERROR PROBABILITY BOUNDS | 190 |
| IV.2.5 ASYMPTOTIC BEHAVIOUR OF VERY LONG CODES | 193 |
| IV.2.6 A GALLAGER TYPE BOUND FOR CODED MODULATION | 196 |
| IV.2.7 CONCLUDING REMARKS | 199 |
| CONCLUSIONS ET PERSPECTIVES | 211 |
| BIBLIOGRAPHIE ET REFERENCES | 219 |

LISTE DES FIGURES ET TABLEAUX.

FIGURES:

Section I.

- Figure 1. Codeur non linéaire de Wei en treillis à 8 états adopté par les normes CCITT pour MODEM à 9600 bps destiné à la transmission MCT/MAQ 32 (efficacité spectrale 4 (bits/s)/Hz) sur le canal téléphonique.....20
- Figure 2. Application bits / symbole pour la modulation codée utilisée par la norme V.32 du CCITT associée au codeur ci-dessus.....21
- Figure 3. Utilisation de la bande téléphonique pour les MODEMS à 19,2 kbits/s.....23
- Figure 4. Illustration de la technique de partition d'Ungerboeck. Partition d'une constellation MDP 8.....31
- Figure 5. Diagramme de treillis à 2, 4 et 8 états pour les codes d'Ungerboeck MCT/MDP.....33
- Figure 6. Codeurs d'Ungerboeck MCT/MDP 8 sous forme systématique avec rétroaction, avec un treillis à 2, 4 et 8 états, respectivement.....34
- Figure 7. Gain potentiel du codage à divers niveaux de probabilité d'erreur.....39
- Figure 8. Capacité et débit critique exprimés en termes de limitation de l'efficacité spectrale.....41
- Figure 9. Constellations MDP et MAQ et leurs applications respectives bits/symbole...49
- Figure 10. Structure générale d'un code de classes latérales $C(\Lambda/\Lambda';C)$67
- Figure 11. Une partition $Z^2/2Z^2$ sur un jeu de signaux à 64 états. L'ordre de la partition est $|Z^2/2Z^2|=4$. Les quatre classes latérales sont représentées.....71
- Figure 12. Code de classe latérale correspondant au code en treillis d'Ungerboeck basé sur un code convolutif ayant un treillis à 4 états, destiné à la transmission de 5 bits/2D et dont le gain fondamental est 3,01 dB.....72

Section II.1

| | |
|---|-----|
| Figure 1. This figure illustrates the iterative construction and bits-to-symbol assignment of a (4M)-point constellation from an M-point one..... | 101 |
| Figure 2. Coordinate systems which are used for both the bits-to-symbol conversion and the fast detection of received data when square constellations and bits-to-symbol assignment of figure 1 are used..... | 102 |
| Figure 3. Generalized cross-constellations with quadrilateral symmetry. Their sizes are 4 (3a), 8 (3b), 12 (3c), 16 (3d), 24 (3e), 28 (3f), and 36 (3g)..... | 102 |
| Figure 4. Bounds on the symbol error probability for a 12-point constellation and the corresponding simulated error rate..... | 104 |
| Figure 5. 7-point generalized constellation with opportunistic channel. The starting (or basic) constellation is the square 4-point one with shaded points. 3 more points are appended to their labels to tell whether the corresponding points belongs or not to the starting constellation..... | 104 |
| Figure 6. Rational logarithm function and its relation with the log function..... | 105 |

Section II.2

| | |
|---|-----|
| Figure 1. Block diagram of a TCM system with differential detection..... | 110 |
| Figure 2. Sensitivity of bit error probability to carrier phase drift for 4-state 8 PSK/TCM Ungerboeck code..... | 110 |
| Figure 3. Block representation of the "differential" trellis code U' associated with an Ungerboeck code U | 112 |
| Figure 4. Trellis diagram for the 2-state U code and its associated 4-state U' code..... | 113 |
| Figure 5. Truth tables for the output Boolean variables Y_n^1, Y_n^0 of the 4-state code.... | 115 |
| Figure 6. 4-state nonlinear convolutional encoder for 8-DPSK/TCM..... | 115 |
| Figure 7. Trellis diagram for the 4-state U code and its associated 8-state U' code..... | 118 |
| Figure 8. Truth tables for the output Boolean variables Y_n^1, Y_n^0 of the 8-state code.... | 118 |
| Figure 9. 8-state nonlinear convolutional encoder for 8-DPSK/TCM..... | 119 |
| Figure 10. Schemes for direct and inverse differential encoders: (a) black box representations, (b) general differential operation, (c) differential encoder for M-PSK..... | 122 |

| | |
|---|-----|
| Figure 11. Simulated error rate vs signal-to-noise ratio (dB). Points correspond to: uncoded 4-DPSK (shaded square) and 4-state TCM/8-DPSK (white square); uncoded 4-PSK (almost shaded square) and 4-state TCM/8-PSK (almost white square)..... | 123 |
|---|-----|

Section III.1

| | |
|---|-----|
| Figure 1. Regions concerning the bounds..... | 140 |
| Figure 2. Sphere packing representation of lattice points..... | 140 |
| Figure 3. Reliability function $E_2(R,C)$ for a Gaussian channel (SNR = 27.1 dB)..... | 140 |

Section III.2

| | |
|--|-----|
| Figure 1. Coherent QAM receiver for lattice-coded signals..... | 159 |
| Figure 2. Lattice figure of merit times the signal-to-noise ratio required for keeping the error rate below 10^{-5} as a function of the dimensionality..... | 159 |
| Figure 3. Putting signal points into a 2-dimensional region with power limitation per dimension. Fig 3.a when only inner points are allowed, fig3.b if points can be chosen inside and on the border of the region..... | 160 |
| Figure 4. Bounds on the performance of lattices at several rates in the presence of additive white Gaussian noise..... | 160 |
| Figure 5. Upper and lower bounds on the block error probability for lattices over an AWGN channel: a) Schläfli lattice D_4 supporting 1.5 (bits/s)/Hz and b) Gosset lattice E_8 supporting 3.0 (bits/s)/Hz..... | 161 |
| Figure 6. Block error rate vs signal-to-noise ratio for the Leech lattice. The data labelled "commercial modem" concerns a 19,200 bits/s modem (7 (bits/s)/Hz, 2,742.86 bauds) manufactured by Motorola, according to [LAN&LON 1989]..... | 162 |
| Figure 7. Pictorial representation of the signal constellation in an Euclidean space \mathbb{R}^n , n very large. Eyesight from a particular signal point s_j . Almost all other signal points lie on a hypersphere of radius $2r$. The picture does not change if the view point becomes the one which represents another signal, for instance s_i | 162 |

Section IV.1

| | |
|--|-----|
| Figure 1. N -dimensional constellation created by taking the Cartesian product of a constituent 2-dimensional constellation with itself $N/2$ times..... | 207 |
| Figure 2. M-point QAM constituent 2-dimensional constellations and their respective 2-dimensional transfer functions..... | 208 |
| Figure 3. Evolution of (normalized) squared Euclidean distance spectrum of random coded constellations based on a 16-QAM constituent constellation for increasing dimensionality N | 209 |

TABLEAUX:

Section I.

| | |
|--|----|
| TABLEAU I. EVOLUTION DES SEMAPHORES..... | 13 |
| TABLEAU II. EVOLUTION DU TELEGRAPHE ELECTRIQUE..... | 14 |
| TABLEAU III. EVOLUTION DES MODEMS SUR LE CANAL TELEPHONIQUE: partie I. MODULATION NUMERIQUE SANS CODAGE..... | 25 |
| TABLEAU IV. EVOLUTION DES MODEMS SUR LE CANAL TELEPHONIQUE: partie II. MODEMS A HAUT DEBIT AVEC CODAGE MODULATION.. | 26 |
| TABLEAU V. GAIN ET NOMBRE D'ETATS DU TREILLIS DE QUELQUES SCHEMAS DE MODULATION CODEE A DEUX DIMENSIONS.... | 35 |
| TABLEAU VI. GAIN POTENTIEL APPORTE PAR LA MODULATION CODEE.... | 38 |
| TABLEAU VII. PARAMETRES DES PRINCIPAUX RESEAUX A FAIBLE NOMBRE DE DIMENSIONS..... | 63 |
| TABLEAU VIII. PARAMETRES DES PRINCIPAUX RESEAUX A FAIBLE NOMBRE DE DIMENSIONS (suite)..... | 63 |
| TABLEAU IX. DENSITE DES RESEAUX Z_2 , E_8 ET A_{24} RAPPORTEE A 2, 8 ET 24 DIMENSIONS..... | 64 |

| | |
|---|----|
| TABLEAU X. DENSITE DE CENTRES DES RESEAUX Z_2 , E_8 ET A_{24} RAPPORTEE A 2, 8 ET 24 DIMENSIONS..... | 64 |
|---|----|

| | |
|--|----|
| TABLEAU XI. GAIN THEORIQUE DU MEILLEUR RESEAU CONNU A PLUSIEURS DIMENSIONS..... | 75 |
|--|----|

Section II.1

| | |
|--|----|
| TABLE I. ENERGY DISTRIBUTION IN THE $2Z^2+(1,1)$ LATTICE..... | 93 |
| TABLE II. PARAMETERS OF SOME GENERALIZED CROSS CONSTELLATIONS..... | 94 |
| TABLE III. ERROR PERFORMANCE OF 32 AND 128-QAM UNCODED SIGNALING SCHEMES..... | 96 |
| TABLE IV. COMPARISON BETWEEN 2D AND MULTIDIMENSIONAL UNCODED SIGNALING SCHEMES..... | 99 |

Section II.2

| | |
|---|-----|
| TABLE I. STATE TRANSITIONS FOR A 4-STATE CONVOLUTIONAL CODE ASSOCIATED WITH THE 2-STATE UNGERBOECK CODE..... | 113 |
| TABLE II. STATE TRANSITIONS FOR THE 4-STATE TCM/DPSK CODE..... | 114 |
| TABLE III. STATE TRANSITIONS FOR A 8-STATE CONVOLUTIONAL CODE ASSOCIATED WITH THE 4-STATE UNGERBOECK CODE..... | 116 |
| TABLE IV. STATE TRANSITIONS FOR THE 8-STATE TCM/DPSK CODE..... | 117 |

Section III.1

| | |
|---|-----|
| TABLE I. RANDOM VARIABLE TRANSFORMATIONS..... | 133 |
|---|-----|

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Voici que près d'un demi-siècle s'est écoulé depuis l'énoncé par Claude Elwood Shannon des théorèmes fondamentaux de la Théorie de l'Information qui ont profondément changé notre compréhension de la nature du processus de communication. Ses nombreuses découvertes scientifiques, cela va sans dire, sont d'une élégance et d'une beauté rares. En proposant une mesure de l'information et une théorie digne de ce nom, il a élevé les télécommunications au rang des sciences.

Il présentait une conception révolutionnaire des parties composant une chaîne de communications, devenue aujourd'hui classique: une source d'information; un émetteur effectuant le codage (de source et de canal) et la modulation; un canal où s'ajoute un bruit aléatoire; un récepteur comprenant le démodulateur et les décodeurs; et finalement un destinataire. Nous nous sommes abstenu d'inclure dans cette thèse la figure correspondante, présente dans 90% des ouvrages sur le codage, mais il y a fallu bien des efforts!

Les principaux aspects étudiés par Shannon vont de la mesure de l'information jusqu'au cryptage (sans parler de l'intelligence artificielle), en passant par le théorème d'échantillonnage, le codage d'une source discrète, le codage de canal, le codage de source avec critère de fidélité, les canaux à plusieurs terminaux. Cette thèse n'est consacrée qu'à une petite partie de cet univers fantastique: le codage de canal, c'est-à-dire le codage en tant qu'outil pour combattre le bruit ajouté pendant la transmission.

Mais comment le bruit affecte-t-il l'information? Les ingénieurs de l'époque préshannonienne croyaient, sur la foi d'un apparent bon sens, que le bruit limitait toujours la qualité de la transmission, ce qui entraînait une passivité fataliste vis-à-vis du bruit. Cela est bien illustré par le passage [HAR 1928] "*... external interference, which can never be entirely eliminated in practice, always reduces the effectiveness of the system*".

Alors que l'on pensait que le bruit dégradait sûrement l'information transmise, Shannon a montré qu'il n'en était rien. Il s'agissait alors d'une idée neuve qui reste aujourd'hui encore surprenante: le sens commun n'associe-t-il pas une transmission erronée à un canal bruyant? Shannon [SHA 1948a, 1948b] a prouvé que le bruit ne limitait que le débit d'information possible, sans imposer toutefois aucune limite à la qualité avec laquelle le message issu de la source peut être restitué au destinataire.

Plus spécifiquement, Shannon démontrait le théorème fondamental du codage de canal [SHA 1949] qui énonçait que le bruit n'empêche pas de réduire autant qu'on le veut la probabilité d'erreur, à condition que le débit d'information reste inférieur à un débit maximal nommé capacité du canal. Sa brillante démonstration est fondée sur un raisonnement, dit de codage aléatoire, qui va être exposé dans la suite. La connaissance de moyens explicites pour obtenir de bons codes faisant défaut, il considérait la famille probabilisée de tous les codes d'un certain type et en calculait la probabilité d'erreur moyenne ou, plutôt, une borne supérieure de celle-ci. L'énoncé du théorème fondamental du codage de canal étant démontré pour cette moyenne, il existait alors au moins un code de la famille qui le vérifiait. Il s'agit cependant d'un théorème d'existence (non de construction). En effet, puisque sa démonstration s'appuie sur des moyennes faites sur des ensembles de codes tirés au hasard, elle ne propose aucune construction particulière permettant de l'atteindre.

En proposant un nouveau regard sur les systèmes de communications, et ayant aussi établi toutes leurs limitations intrinsèques, Shannon ouvrait de multiples horizons pour les recherches, donnant naissance en particulier à une nouvelle classe de chercheurs, celle des "théoriciens du codage".

La clef du codage fut clairement exhibée par Shannon: il est plus avantageux d'attendre d'avoir plusieurs points puis de les décoder ensemble en les considérant comme un point d'un espace ayant un grand nombre de dimensions que de les décoder au fur et à mesure de leur arrivée.

La probabilité d'une transmission sans erreur étant établie, les moyens de l'atteindre par une construction explicite devenaient objet de recherche. Les études de codage ont donc abouti à tout un arsenal de moyens et se sont avérées d'une fécondité surprenante [BER 1974]. La première classe de codes découverte, celle de Hamming, a même été citée par Shannon à titre d'exemple [e.g., WEA&SHA 1949]. Au début des années 40, les *Bell Telephone Lab.* utilisaient déjà un code détecteur d'erreur similaire au code de parité (5,4) dans deux ordinateurs à relais [THO 1983]. Au début des années 50, Richard Hamming, en tant que chercheur chez Bell, a été amené à se poser la question: "*Damn it, if the machine can detect an error, why can't it locate the position of the error and correct it?*" [THO 1983]. Il créa alors les premiers codes binaires capables de corriger automatiquement des erreurs [ibidem]. Presque simultanément, Marcel Golay découvrait les seuls autres codes parfaits existants. D'autre part, David Slepian exhibait et exploitait, en 1957, la structure linéaire de code en bloc. Les moyens explicites de codage commençaient leur longue carrière.

Les débuts des années 60 furent profondément marqués par le développement de très puissantes techniques algébriques de codage: les codes BCH, les codes de Reed-Solomon (notés ici codes RS [REE&SOL 1960]) ainsi que les codes de résidu quadratique. La fin de cette période coïncide avec l'apparition de l'algorithme de décodage de Berlekamp [voir BER 1984]. Les années 70 ont vu l'invention des codes de Goppa, des codes alternants et de bien d'autres généralisations [BER 1984].

Dans le cadre du codage par correction directe des erreurs (FEC), deux stratégies sont utilisées: d'une part, les codes en bloc avec de fortes propriétés algébriques, et d'autre part les codes convolutifs (pour un point de vue intéressant à ce propos, voir [BAT 1984]).

La classe des codes convolutifs fut créée en 1954 par Peter Elias (au MIT). Le décodage séquentiel, pertinent à cette classe de codes, s'est développé fortement au début des années 60, en particulier avec les contributions de Wozencraft, Fano et Jelinek. Cependant, il a fallu attendre jusqu'en 1967 pour que l'outil le plus puissant de la théorie et de la pratique du codage soit découvert: l'algorithme de Viterbi, le plus remarquable développement du codage depuis sa création [VIT 1967]. D'autre part, l'utilisation des sondes spatiales a fourni une formidable incitation au développement du codage de canal.

Du point de vue du décodage, l'exploitation de propriétés algébriques exige des opérations sur les symboles de l'alphabet discret du code. Cela implique une décision préalable sur les signaux reçus, qui remplace leur position vraie dans l'espace des signaux par la donnée du point de la grille utilisée au codage qui en est le plus proche. Cette décision, dite "ferme", implique la perte de la marge qui renseigne sur sa fiabilité. De meilleures performances sont obtenues avec des techniques qui utilisent l'information probabiliste reçue et que l'on désigne par décodage souple ou à décision pondérée. L'avènement de la pondération dans les algorithmes de décodage constitue une importante étape dans les progrès du codage. En France, Battail a été le plus important défenseur et propagateur de cette idée. L'adaptation du décodage en treillis (dite du type Viterbi) aux codes en bloc a été développée principalement par Battail et Wolf [BAT&DEC 1976, BAH *et al.* 1974, WOL 1978, BAT 1983].

La fonction du codage peut être interprétée de deux façons distinctes. Si l'on suppose que le rapport signal à bruit est fixé, une augmentation du débit entraîne une dégradation des performances. Le codage apparaît alors comme un moyen d'améliorer le débit sans dégrader la qualité de transmission. Si l'on admet que c'est le taux d'erreur souhaité que l'on fixe, une augmentation du débit entraîne un besoin accru de rapport signal à bruit. Voilà le codage envisagé comme moyen d'épargner de la puissance de

l'émetteur.

• LE CODAGE EFFICACE AVEC LIMITATION DE BANDE.

Etant donné un canal de bande passante B , on voit aisément que la seule façon d'augmenter son débit d'information est d'en accroître l'efficacité spectrale. Cela entraîne l'emploi de constellations à grand nombre d'états, c'est-à-dire une augmentation de la taille de l'alphabet des symboles. En admettant un rapport signal-à-bruit fixé (car le canal est donné), il en résulte naturellement une dégradation de la performance du système, exprimée par sa probabilité d'erreur. En effet, les différents signaux deviennent de moins en moins discernables. Dans ce cas, la nécessité du codage apparaît clairement pour améliorer la performance en réduisant le taux d'erreur à des niveaux tolérables.

Lorsqu'il est fait allusion à l'introduction de redondance, un premier défi se pose. Pour une taille donnée de l'alphabet, la redondance ne peut être introduite qu'en augmentant le nombre des symboles émis. Cela se traduit soit par une perte en débit utile, soit par une augmentation de la bande de fréquences exigée pour la transmission du signal.

La première constatation qui s'impose est hélas qu'on ne peut jamais s'approcher de la capacité du canal en utilisant les techniques classiques. En effet, l'augmentation de la largeur de bande n'est pas permise dans des canaux à bande étroite. Par ailleurs, même si l'on accepte un élargissement de la bande, ceci revient à changer le canal et donc sa capacité! La "réduction du débit utile de transmission" pour loger la redondance n'est guère attrayante et s'avère paradoxale, car on souhaite l'augmenter. Ainsi, dans la majorité des applications pratiques, le gain de codage n'était obtenu qu'au prix d'un élargissement du spectre, ce qui soulevait la question de la possibilité d'appliquer le codage aux canaux à bande étroite et éloignait le codage du paradis shannonien.

Cette contradiction inhérente au processus classique de codage a longtemps différé l'application des codes correcteurs d'erreurs aux canaux à bande étroite. A cela s'ajoute un argument développé par Forney [FOR 1970] selon lequel on parvenait, sans codage, à s'approcher de la capacité au fur et à mesure que l'on pénétrait dans la région limitée en bande passante. Cela explique en partie l'inexistence de systèmes codés opérant sur le circuit téléphonique (par exemple dans les MODEMS).

D'autre part, les canaux de satellites sont bien modélisés par du bruit blanc gaussien avec un rapport signal à bruit assez faible (typiquement quelques décibels) et une bande passante assez large, où la puissance coûte quelques centaines de milliers de dollars par décibel. La première application du codage de canal à grande échelle a été le code d'Odenwalder, un code convolutif (2,1) de longueur de contrainte 7, sur des canaux de satellite [e.g., WU *et al.* 1987]. Cela n'est pas surprenant dans la mesure où cet ensemble de caractéristiques est très favorable à l'application du codage. En 1982, des codes de Reed-Solomon furent incorporés dans les normes des télécommunications spatiales [YUE 1990]. La limitation de largeur de bande ne se faisant pas sentir à l'époque, ces canaux furent considérés comme *limités en puissance*. Ainsi, ce n'est pas par hasard que les techniques classiques de codage n'ont eu de réussite remarquable que dans le domaine des télécommunications spatiales et par satellite. Les sceptiques à propos du codage affirmaient alors "*Coding is dead except for the deep-space channel*" [MAS 1974]. A présent, même ces canaux sont considérés comme *limités en bande passante* et les nouvelles techniques dont traite cette thèse sont préférables [HAG 1989, YUE *et al.* 1990].

Ainsi, pendant longtemps, le codage est apparu comme un palliatif, une technique pour faire passer la qualité des données de problématique à acceptable, le prix à payer étant une réduction du débit utile ou une augmentation de bande passante. Or, Shannon avait démontré que le seul prix à payer pour réduire le taux d'erreur est un accroissement de la complexité du décodeur, pourvu que le débit utilisé soit inférieur à la capacité du canal.

Pourrait-on s'affranchir de ces difficultés? La réponse n'est venue que beaucoup plus tard. Il fallait donc développer des techniques permettant d'améliorer les performances d'une transmission numérique sans exiger ni une plus grande largeur de bande passante, ni un sacrifice sur le taux d'émission de l'information utile (cela paraissait plutôt impossible aux yeux des ingénieurs de l'époque).

La nouvelle révolution ne s'est produite qu'à la fin des années soixante-dix / débuts des années quatre-vingts grâce à l'ingénieuse idée de Gottfried Ungerboeck [UNG 1976] qui a permis de sortir de cette impasse et a complètement changé le *statu quo* du codage. Le premier pas dans cette direction a consisté à considérer des techniques de codage où la redondance est introduite par en élargissant le jeu de signaux utilisé au lieu d'ajouter des symboles de parité. Cette façon de procéder, où la redondance est obtenue

par l'emploi d'un espace des signaux plus grand que celui qui est nécessaire à la modulation non codée, est aujourd'hui appelé codage avec expansion de l'alphabet. La deuxième étape a consisté à réaliser conjointement la conception du code et celle de la modulation. Cela est devenu réalisable à l'aide d'une technique développée par Ungerboeck et nommée "partition des ensembles" (*set partitioning*). Une démonstration rigoureuse que les codes d'Ungerboeck n'entraînent pas d'élargissement de la bande a été présentée par Biglieri [BIG 1986].

Bien entendu, pour gagner en débit sans dégrader les performances, il faut payer un prix. Car, fort insidieusement, il se trouve que plus la façon de coder s'approche de l'idéal, plus les délais deviennent longs lors du processus de décodage. Les promesses ultimes de la théorie ne sont atteints qu'asymptotiquement en construisant des jeux de signaux (codes) dans des espaces à nombre de dimensions très élevé. Le but de cette thèse est principalement de contribuer à rendre plus compréhensible le mécanisme du codage avec des constellations de signaux de dimension finie et de fournir quelques réflexions géométriques à propos des phénomènes particuliers aux espaces euclidiens à nombre de dimensions très élevé.

Les transmissions de données ont connu une véritable explosion, particulièrement due à l'usage intensif des ordinateurs qui caractérise l'époque actuelle. Par conséquent, les besoins en débit d'information se sont accrus. Le codage de canal s'avère maintenant indispensable. Néanmoins, il n'est pas facile de s'approcher de la capacité, objectif très difficile à atteindre à cause de la complexité intrinsèque des algorithmes de décodage [e.g., BER *et al.* 1978].

L'emploi pratique de systèmes de codage élaborés devient de jour en jour plus attrayant du point de vue économique et cette tendance paraît destiner à continuer. On peut, en parodiant J.K. WOLF, affirmer "*People who reach(ed) the conclusion that coding is (was) dead are (were) dead wrong*".

TRAVAUX ANTERIEURS

"It is better do the right problem the wrong way
than do the wrong problem right way" R.W. HAMMING.

CHAPITRE I

INTRODUCTION AUX SYSTEMES DE MODULATION CODEE

I.1. BREFS RAPPELS SUR L'HISTOIRE DES COMMUNICATIONS.

I.1.1 ANCIENNE HISTOIRE DES TRANSMISSIONS NUMERIQUES.

On présente tout d'abord quelques notes historiques dont le but est de décrire brièvement l'évolution des systèmes de télécommunications numériques en exhibant d'anciennes réalisations parmi les plus importantes qui concernent la théorie des communications numériques. L'aperçu est à la fois trop rapide et incomplet.

En 1623, Francis Bacon constatait que l'homme pouvait communiquer à distance au moyen de simples signaux binaires. En 1771, Chappe expérimentait un sémaphore (dit télégraphe visuel) entre les municipalités de Parce et Brulon. Son système utilisait initialement un alphabet de 26 signaux, qui a ensuite été porté à 66 symboles, dont les lettres majuscules, minuscules et quelques chiffres. Son développement [cf. STU 1984] est présenté dans le tableau I [ANDR 1989].

TABLEAU I. EVOLUTION DES SEMAPHORES.

| année | entrepreneur | liaison (localités) |
|---------------|-----------------------|--|
| 1794 | Carnot | Paris - Lille |
| 1795 | Lord Murray | Londres - Portsmouth, Plymouth, Liverpool |
| 1795- 1845 | Napoléon | Toute la France |
| 1835 | Tsar Nicolas I | Moscou - Petrograd |
| 1835 | Frédéric IV de Prusse | Berlin - Coblenze |

En 1809, Samuel Thomas Soemmering a construit en Bavière le premier télégraphe électromécanique. Le développement du télégraphe électrique est esquissé dans le tableau II ci-dessous. Déjà en 1845, des connections télégraphiques étaient disponibles entre Paris et Rouen, ainsi que Londres et Porthmouth, par exemple. Le télégraphe électrique fut ensuite sensiblement amélioré par W. Siemens en Allemagne.

TABLEAU II. EVOLUTION DU TELEGRAPHE ELECTRIQUE.

| | | |
|-------------------|-------------------------------------|------------|
| 1832-1835 | Paul Schilling Von Canstatt | Bonn |
| 1833 jusqu'à 1838 | C. Friedrich Gauss et W. Weber | Göttingen |
| 1836 | William Cooke et Charles Wheatstone | Londres |
| 1837 | Samuel F.B. Morse | Etats Unis |

En 1851, la France, la Belgique et la Prusse signaient un pacte sur le trafic télégraphique, auquel plusieurs pays ont ensuite adhéré. Entre 1850 et 1851, un câble maritime était lancé entre Douvres et Calais. En 1865, l'Union Télégraphique Internationale - UTI a été créée à Paris [BEL *et al.* 1984]. Elle adopta, en 1903, les appareils du type Baudot (un code avec 5 impulsions par lettre, créé en 1874). Le premier câble transocéanique était lancé par l'américain Cyrus W. Field en 1866 après de nombreuses péripéties [ZWE 1990].

Entre 1895 et 1901, Guglielmo Marconi (prix Nobel 1909) a perfectionné son télégraphe sans fil. Les télécommunications numériques changeaient définitivement de visage. La première communication radioélectrique transocéanique par satellite a eu lieu entre Holmdel, New Jersey, et la station française de Pleumeur-Bodou, en 1962 [IEEE Comm. Mag, May 1984, p.23].

L'idée d'une mesure quantitative de l'information fut avancée en 1918 par Hartley. Il définissait la quantité d'information comme égale au logarithme du nombre des séquences de symboles permises [HAR 1918]. En 1928, Nyquist a montré que la transmission de signaux télégraphiques à un débit donné exige une certaine bande passante [NYQ 1928]. En termes plus précis, la bande spectrale minimale nécessaire à une interprétation non ambiguë est égale au débit numérique en symboles/seconde (bauds) et ne dépend pas de la taille de l'alphabet de symboles (nombre de niveaux d'amplitude employé). Ce résultat, étroitement lié au théorème d'échantillonnage, a été formellement démontré par Shannon [OLI *et al.* 1948, SHA 1949]. En 1920, G. Valensi (France) a développé la première suggestion sérieuse pour le multiplexage temporel de canaux téléphoniques. Il a offert son idée aux Etats-Unis en reconnaissance de leur aide lors de la 1^{ère} guerre mondiale [MIL 1984]. En 1937, l'anglais A.H. Reeves concevait, aux laboratoires de l'ITT à Paris, l'idée de la modulation d'impulsion en amplitude (MIA). En 1945, Mr. Deloraine inventait en France le multiplexage temporel (MDT) [STU 1984].

Les modulation numériques ont évolué: les plus anciens systèmes ont utilisé la modulation par déplacement de fréquence (MDF) et la modulation en phase (MDP). Ensuite Campopiano et Glazer ont introduit une technique combinant la modulation en amplitude et en phase pour la communication numérique [CAM&GLA 1962], développant les modulations appelées aujourd'hui Modulation d'Amplitude en Quadrature (MAQ).

De nombreuses approches ont été utilisées pour développer des constellations à deux dimensions. Le problème d'optimiser des jeux de signaux à deux dimensions avec des composantes en phase et en quadrature (du genre MAQ) sous une contrainte de puissance a été abordé en utilisant une procédure de recherche du gradient minimal. Les constellations optimales à deux dimensions ainsi obtenues ont apporté des gains d'énergie moyenne de l'ordre de 0,5 dB par rapport aux modulation MAQ classiques [FOS *et al.* 1974]. Les performances de la modulation numérique en Amplitude et Phase avec un alphabet M -aire ont été étudiées en supposant une détection cohérente dans un canal à bruit gaussien. Diverses configurations de constellations (de taille M allant de 4 à 128 points) ont été étudiées en ce qui concerne la probabilité d'erreur par symbole et le rapport signal à bruit, moyen et de crête. Les dégradations causées par les non-linéarités après le passage du signal modulé dans un amplificateur du type Tube à Ondes Progressives (TOP) furent aussi examinées [THOM *et al.* 1974]. Cependant, les faibles gains apportés par l'optimisation de la modulation indiquaient que l'avenir des transmissions numériques était lié au codage.

I.1.2 L'IDEE ET L'HISTOIRE DE LA MODULATION CODEE

"The recognition that coding is useful on bandwidth-limited channels is a recent phenomenon and stems from the remarkable successes achieved by Ungerboeck, whose pioneering paper ... where it opened the eyes of many engineers to new possibilities for coding" J. MASSEY.

La théorie de la modulation codée est un domaine fort fascinant, résultat d'une "alchimie" entre la théorie du codage correcteur d'erreurs et des techniques de modulation numérique. Cette hybridation s'est avérée d'une fécondité surprenante et l'on observe aujourd'hui une véritable explosion d'intérêt à propos de ce sujet. La clef de son succès réside dans le projet conjoint d'un code lié à une modulation de façon à bien les adapter au bruit du canal. Son invention remonte à Ungerboeck après une épopée remarquable que l'on va maintenant résumer.

La modulation codée en treillis (MCT) a été pour l'essentiel conçue par Ungerboeck en 1976 [UNG&CSA 1976]. Son article original, proposé aux *IEEE Trans. Info. Theory*, n'a pas été immédiatement accepté. Cependant, quelques temps après, l'apport des nouveaux concepts est devenu de plus en plus évident, et les éditeurs ont reconsidéré leur décision initiale! [CAL 1990a]. En conséquence, l'article d'Ungerboeck a été enfin publié en 1982 [UNG 1982], gagnant par la suite le prix du meilleur article (*best paper award*) du groupe de théorie de l'information de l'IEEE. Il ne fait aucun doute que toute la conception des systèmes numériques codés a connu alors une véritable révolution.

Le premier défaut du codage classique, tel que nous l'avons décrit dans l'introduction générale, découlait de la façon inadéquate d'apporter la redondance, interdisant une utilisation efficace de la capacité offerte par le canal. D'autre part, la difficulté à trouver des moyens de codage efficaces venait aussi du fait que leur construction était basée sur des codes, principalement algébriques, conçus indépendamment de la modulation. On mentionne que l'idée que la maximisation de la distance libre de Hamming n'était pas un bon critère pour le projet de codes performants destinés au canal gaussien était plus ou moins répandue à l'époque. Néanmoins, le codage continuait à se développer (comme à son origine) en tant que mathématique discrète, s'éloignant de sa raison d'être: le canal.

Massey fut un des pionniers à reconnaître que le codage et la modulation devraient être envisagés dans leur ensemble et non comme des entités isolés [MAS 1972]. Ne se contentant pas de reprocher verbalement le pessimisme quant à l'avenir du codage, il proposa une autre direction pour aborder le problème (*vox clamantis in deserto*). Ungerboeck, étant aussi à Zurich (chez IBM), a pris connaissance de ces idées et probablement a saisi leur importance. Par la suite, il a fait un séjour sabbatique au Centre d'Etudes et Recherches IBM, La Gaude, France. Là, ses intérêts de recherche se sont portés vers le codage de canal [WOL&UNG 1986,p.773] et plus particulièrement vers son rapport étroit avec la modulation .

A cet égard Digeon, qui travaillait à l'IBM à la même époque, remarqua [DIG 1976] *"A bande passante du canal constante, à vitesse de modulation constante, à rapport signal sur bruit en puissance moyenne constant, nous pouvons envisager d'améliorer la probabilité d'erreur de la façon suivante: si nous doublons le nombre de signaux de notre modulation, sous les hypothèses précédentes ..."* (sic).

Dès les premiers résultats basés sur un système codé à huit niveaux d'amplitude, comparé à une modulation non codée à quatre niveaux d'amplitude [DIG 1976, DIG 1977], il trouva des gains de codage significatifs. Cela montrait définitivement que, contrairement à ce que l'on avait cru, le gain apporté par le codage pouvait compenser la perte causée par l'élargissement de l'alphabet. Il rejoignait l'idée de Massey que la recherche d'un code ayant de bonnes caractéristiques doit se faire en liaison avec l'étude de la modulation.

Ungerboeck s'est donc proposé d'utiliser la technique de codage où la redondance est introduite par l'élargissement du jeu de signaux (c'est-à-dire l'extension de l'alphabet de la constellation) au lieu d'ajouter des symboles/bits de parité, afin d'éviter d'augmenter la largeur de bande. Cette étape franchie, la difficulté suivante consistait alors à choisir un code-modulation (selon les idées de Massey, par exemple) bien adapté au bruit. Comment le pouvait-on ? Dans le cas de bruit blanc gaussien, il s'agissait de construire des codes garantissant une bonne distance (libre) euclidienne plutôt que de Hamming. Ici apparaît la contribution magistrale d'Ungerboeck: la technique de partition des ensembles décrite dans la prochaine section.

L'importance du formidable travail d'Ungerboeck a été telle qu'il a créé un nouveau sujet de recherches en théorie de l'information: la modulation codée, objet de cette thèse. Néanmoins, il conviendrait de remarquer qu'il y a deux approches différentes au problème de la modulation efficace en bande passante combinée au codage. D'abord, la "Modulation à Phase Continue" (MPC) développée par Sundberg, Aulin et Anderson

[SUN 1986, WIL 1986], ensuite la "Modulation Codée en Treillis" d'Ungerboeck. La première approche consiste en une classe de schémas de modulation numérique à amplitude constante où le codage est conçu conjointement avec la modulation dans le but de maximiser la distance euclidienne minimale. On impose la continuité aux variations de la phase du signal (support de l'information) et des spectres assez étroits découlent de cette contrainte. En outre, l'amplitude constante constitue un avantage pour la transmission dans des canaux à forte non-linéarité. Quoique une vaste littérature ait été consacrée à ce sujet [AND *et al.* 1986], il ne sera pas abordé dans cette étude. Cela pourrait en soi faire l'objet d'une thèse. D'ailleurs, le codage MCT s'est avéré beaucoup plus performant et moins complexe que les MPC, et il devient même de plus en plus évident qu'il est supérieur dans la plupart des applications pratiques du codage de canal.

En outre, il faut aussi accorder de l'importance à un autre point acquis définitivement par le travail d'Ungerboeck. Il s'agit de comparer différents systèmes dans des conditions de même débit d'information et bande passante, autrement dit, des systèmes de même efficacité spectrale. Cela permet une comparaison nette et équitable pour évaluer le gain du codage.

La modulation codée peut être mise en œuvre à l'aide des codes convolutifs, ce qui a pour résultat la modulation codée en treillis (MCT) ou bien par l'intermédiaire des codes en bloc, ce qui est connu comme modulation codée en bloc (MCB). Ayant ainsi esquissé quelques détails sur sa création, nous allons consacrer la prochaine section à décrire brièvement la méthode utilisée par Ungerboeck.

I.1.3 L'EVOLUTION DES MODEMS.

De toute évidence, une demande croissante pour des débits numériques élevés se manifeste dans l'évolution de la téléinformatique. L'importance de la transmission de données par voie téléphonique augmente jour après jour et le besoin de codage se fait sentir d'autant plus que le débit s'accroît. Les techniques étudiées dans cette thèse sont devenues fondamentales dans ce domaine, et se sont montrées de très loin supérieures au codage classique. Bien entendu, l'augmentation du débit peut permettre des économies sensibles sur le nombre de lignes louées ainsi que la création de nouvelles applications. Le canal téléphonique présente une largeur de bande grossièrement comprise entre 300 et 3400 Hz avec un rapport porteur à bruit de l'ordre de 20 à 30 dB (typiquement 26 dB).

Les premiers MODEMS ont été conçus chez BELL entre la moitié et la fin des années 50. Ils étaient à très faible débit, 300 et 600 bits/s, et la modulation numérique utilisée était la modulation binaire par déplacement de fréquence MDF.

Pendant la décennie 60, de nouveaux modems à débit numérique moyen (1200, 2400 et 4400 bits/s) furent développés avec des modulations de phase MDP. Le succès commercial des MODEMS à haut débit ($\geq 9\ 600$ bps) a commencé en 1970 avec le CODEX 9 600, utilisant une modulation MAQ. Les chercheurs ont concentré leurs efforts pour trouver de nouvelles techniques augmentant le débit et améliorant la fiabilité des transmissions numériques. Plusieurs entreprises ont donc tenté, sans aucun succès, d'incorporer les techniques classiques du codage correcteur d'erreur dans le projet de MODEMS en vue d'augmenter le débit sans une dégradation excessive de performance.

On a cru jusqu'à la fin de cette décennie que des débits numériques plus élevés ne seraient jamais possibles dans des systèmes commerciaux. A ce propos Holsinger [HOL 1977] remarque "*Sad but true for engineers of the world, the days of dramatic speed and performance breakthroughs were over almost as soon as they began*" (sic). Le discours des experts pourrait être résumé ainsi: "De toute évidence, le canal téléphonique impose de sévères limitations aux performances des modems; en pratique, un débit de 9600 bits/s apparaît comme la limite ultime pour une transmission fiable de données dans les voies disponibles" [HOL 1977, NYM *et al.* 1982].

Au début des années 80, la première génération de MODEMS à haut débit, à 14400 bits/s, est apparue chez PARADYNE, conçue à partir d'une modulation MAQ 64 non codée sur lignes spécialisées. Elle a été suivie, en 1981, par le CODEX 14 400 avec une constellation bidimensionnelle hexagonale [FOR *et al.* 1984]. Quoi qu'il en soit, les performances atteintes sur le canal téléphonique n'étaient pas assez bonnes pour justifier l'idée que des débits supérieurs à 9 600 bits/s seraient largement introduits sur le marché.

Un très lucide rappel sur l'introduction du codage dans les modems sur le canal téléphonique est présenté par Ungerboeck [UNG 1987a,1987b]. Les premiers MODEMS commerciaux à MAQ codée en treillis sont apparus en 1984, fonctionnant au débit de 14400 bits/s. Maintenant, un code non linéaire en treillis à 8 états [WEI 1984b] avec un gain de 4 dB a été adopté dans les normes du CCITT pour les modems à 9 600 bps sur réseau téléphonique public à 2 fils (14,4 kbps sur des lignes spécialisées à 4 fils). La non-linéarité est introduite pour permettre une protection contre les ambiguïtés de phase de 90° [norme V.32, CCITT 1989]. Ce codage-modulation est illustré par les figures 1 et 2. L'application de la modulation codée aux canaux téléphoniques ne se résume pas aux seuls codes en treillis. En 1986, un MODEM à 14 400 bits/s sur des lignes spécialisées basé sur le codage de réseau (le réseau de Gosset E_8) a été commercialisé par MOTOROLA.

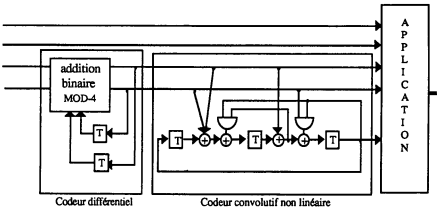


Figure 1. Codeur non-linéaire de Wei en treillis à 8 états adopté par les normes du CCITT pour un MODEM à 9600 bps destiné à la transmission MCT / MAQ 32 (efficacité spectrale 4 (bits/s)/Hz) sur le canal téléphonique.

• 2^{ème} étape:

Les premiers modems commerciaux sont apparus chez Bell au débit numérique de 300 bps, puis de 1200 bps. Une modulation par déplacement de fréquence MDF était utilisée. De nouvelles techniques de modulation ayant un rendement spectral supérieur furent introduites pour répondre à la demande accrue de transmissions numériques, ainsi que de nouvelles techniques de synchronisation de porteuse et de rythme. L'égalisation, d'abord manuelle puis adaptative, a été introduite. Tout cela a permis d'élargir la bande de fréquences utilisée de 1200 Hz à 1600 (puis 2400) Hz.

• 3^{ème} étape:

Cette étape est surtout marquée par l'apparition de l'égalisation fractionnée, des annuleurs d'écho et l'utilisation de processeurs de traitement du signal [STE 87, BIN 88]. Les développements technologiques ont abouti à une réduction considérable de la taille des équipements. Au cours des années 70, plusieurs entreprises ont tenté, sans succès, d'incorporer des techniques classiques de codage correcteur d'erreur dans le projet de MODEMS. Plusieurs MODEMS à 9600 bps ont été développés avec un égaliseur auto-adaptatif, des annuleurs d'écho puissants et même du codage convolutif traditionnel avec décodage basé sur l'algorithme de Viterbi. Le gain atteint par le codage dissocié de la modulation n'était que d'environ 1 à 2 dB. Le MODEM est devenu un outil banal dans la communication de données, ayant pour principaux fabricants: Intertel, IBM, General Datacomm, Codex Cooperation, Paradyne, Bell Systems, MOTOROLA etc.

La transition de cette période vers la suivante est marquée par PARADYNE. La qualité des lignes téléphoniques ayant été aussi sensiblement améliorée, une transmission à 14 400 bits/s a été réalisée avec la MAQ 64.

• 4^{ème} étape:

En 1985, un MODEM à 19 200 bits/s a été produit chez NEC [HIR 1985]. L'idée utilisée y consistait essentiellement à utiliser en parallèle plusieurs sous-canaux orthogonaux dans les 3 kHz de bande du canal téléphonique, système dit multiporteuse [HIR 1980, HIR 1981]. Pour un grand nombre de tels sous-canaux, l'égalisation d'amplitude et de phase dans chacun d'eux devient très simple, puisque la fonction de transfert y est avec une bonne approximation une fonction affine de la fréquence. Cependant, ces MODEMS exigent des circuits de synchronisation et des techniques de récupération de phase assez élaborées et utilisent un grand nombre de porteuses pilotes

pour composer les signaux de synchronisation.

Des MODEMS à 19 200 bits/s ayant un plein succès commercial ne furent fabriqués qu'après la parution du sensationnel article de Wei sur la MCT multidimensionnelle, qui a permis la mise en œuvre de codes du type d'Ungerboeck avec une petite expansion de l'alphabet [WEI 1987]. On décrit brièvement dans la suite quelques détails techniques d'un tel MODEM sur lignes spécialisées à 4 fils [cf. WEI 1987 p.494, TAN *et al.* 1987, BAU&MIT 1989]. Il met en œuvre un codage en treillis très perfectionné. Le système utilise une modulation de phase et d'amplitude d'une porteuse à 1 850 Hz et un égaliseur fractionnaire automatique 3T/4 pour compenser les distorsions provoquées par la ligne en bord de bande. La modulation, dite MAQ généralisée, emploie une constellation en croix de 160 points. Le filtrage est réalisé par un filtre en cosinus surélevé à coefficient d'arrondi très petit, de l'ordre de 0,09375. La fréquence de modulation n'a pas la valeur traditionnelle de 2 400 Hz: elle vaut 2 742,86 Hz (2 400 multiplié par 8/7).

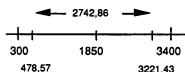


Figure 3. Utilisation de la bande téléphonique pour les MODEMS à 19,2 kbits/s.

Un signal codé est engendré soit à l'aide d'un codeur convolutif de taux 3/4 ayant un treillis à 64 états [WEI 1987,p.494], soit par un codeur en treillis à 16 états [p.ex, UNG 1987b,p.21]. On rappelle que le codeur travaille à la cadence de 685,71 Hz, soit le quart du rythme de la modulation. Le décodage pondéré est fait par l'algorithme de Viterbi. Le gain asymptotique atteint par rapport à la MAQ 128 est de 5,41 dB (5,27 dB selon [FOR 1988,Tableau 7,p.1144]) alors que le gain effectif est de 4,63 dB (4,49 dB). Le premier MODEM produit chez CODEX utilisait le code MCT de Wei, à huit dimensions et 64 états [PAH&HOL 1988,p.23], alors que d'autres modems récemment apparus ont adopté un treillis à 16 états [TAN *et al.* 1987, BAU&MIT 1989].

Récemment, un MODEM à 19 200 bits/s est apparu chez MOTOROLA où un codage à 24 dimensions est réalisé par le réseau de Leech [LAN&LON 1989]. Des blocs successifs de 96 bits de données sont codés en des points de l'espace de signaux (1 bloc étant représenté par un 1 point à 24 dimensions). Les symboles associés à un point à 24

dimensions appartiennent à un répertoire de $1\,024 = 2^{10}$ points pour 2 dimensions. Le nombre total des signaux possible dans l'espace euclidien à 24 dimension est donc $(2^{10})^{12} = 2^{120}$ dont seulement 2^{96} appartiennent au réseau de Leech. Le débit d'information est donc de 96 bits / 24 dimensions = 8 bits/2D ou 8 bits/T. Le rythme étant 2 400 bauds, un symbole MAQ à deux dimensions est émis toutes les $T = 1/2400$ secondes et un signal à 24 dimensions (point du réseau de Leech) toutes les $T_A = N/2 T = 12 \cdot 1/2400 = 1/200$ seconde. Pour donner une idée de la complexité inhérente à un nombre de dimensions aussi élevé, il faudrait plus de 10^{19} années pour transmettre la liste exhaustive des mots du code, correspondant à chaque point du réseau, sans répétition !

On constate cependant que l'utilisation d'une bande d'à peine 2 400 Hz exige une efficacité spectrale de 8 (bits/s)/Hz pour atteindre le débit proposé. Pour ce type de codage, étant donné que le facteur d'expansion de l'alphabet est de 4, il faudrait employer une constellation MAQ à 1024 états! Dans les MODEMS à 19,2 kbps actuellement disponibles, on adopte plutôt une bande de 2 742,86 Hz pour le canal téléphonique. La constellation bidimensionnelle constituante MAQ, nécessaire pour construire le réseau de Leech, doit tout même avoir 512 points. Ce qui précède semble indiquer que des MODEMS basés sur le réseau de Leech Λ_{24} sont beaucoup moins attrayants pour les canaux réels que ceux construits selon une MCT multidimensionnelle, vu leur besoin accru d'expansion de l'alphabet (voir III.2.4.).

Les tableaux suivants présentent un aperçu des principales améliorations et développements des modems sur le canal téléphonique, d'abord sans codage, puis en utilisant plusieurs techniques de modulation codée déjà décrites.

TABLEAU III. EVOLUTION DES MODEMS SUR CANAL TELEPHONIQUE.

partie I. MODULATION NUMERIQUE SANS CODAGE.

| Modulation | efficacité spectrale | bande passante | débit numérique | MODEM | remarques |
|------------|-------------------------|-------------------|--------------------|----------|-----------------------------|
| MDF 2 | 1 bit/s/Hz | 300 Hz | 300 bps | Bell 103 | années 50 |
| MDP 2 | 1 bit/s/Hz | 1200 Hz | 1200 bps | Bell 202 | fin des années 50 |
| MDP 4 | 2 bits/s/Hz | 1200 Hz | 2400 bps | Bell 201 | 1962 |
| MDP 8 | 3 bits/s/Hz | 1600 Hz | 4800 bps | Milgo | 1967 égalisation manuel. |
| MAQ 16 | 4 bits/s/Hz | 2400 Hz | 9600 bps | Codex | 1971 égalisation adapt. |
| MAQ 64 | 6 bits/s/Hz | 2400 Hz | 14400 bps | Paradyne | 1980 |

TABLEAU IV. EVOLUTION DES MODEMS SUR CANAL TELEPHONIQUE.

partie II. MODEMS A HAUT DEBIT AVEC CODAGE MODULATION.

| Modulation constituante | efficacité spectrale | bande passante | débit numérique | MODEM | Codage |
|----------------------------|-------------------------|-------------------|--------------------|----------|---|
| MAQ 32 | 4 bits/s/Hz | 2400 Hz | 9600 bps | AT&T | MCT à 4 états 1984 |
| MAQ 128 | 6 bits/s/Hz | 2400 Hz | 14400 bps | divers | MCT à 8 états 1984 |
| MAQ 128 | 6 bits/s/Hz | 2400 Hz | 14400 bps | MOTOROLA | Réseau Eg 1986 |
| MAQ 256 | 7 bits/s/Hz | 2400 Hz | 16800 bps | CODEX | MCT à 8 états 1984 |
| O-MAQ (200 Hz) | 7 bits/s/Hz | 2800 Hz | 19200 bps | NEC | Multiporteuse 14 sous-canaux 1985 |
| MAQ 160 | 7 bits/s/Hz | 2743,86 Hz | 19200 bps | CODEX | MCT 8-dimension (à 64 états) 1987 |
| MAQ 512 | 7 bits/s/Hz | 2743,86 Hz | 19200 bps | MOTOROLA | Réseau A ₂₄ 1988 |

I.1.4. DE LA POSSIBILITE DE TROUVER DES BONS CODES.

"In brief, this would result from the many symmetries of the MDS codes, which mimic the average regularity of randomness"
G. BATTAIL.

On sait qu'un code aléatoire long est un très bon code dans le sens que la probabilité d'erreur devient très petite au fur et à mesure que sa taille augmente. Ainsi l'écart entre la distribution de poids (ou de distance) du code aléatoire et d'un code donné semble bien être un critère raisonnable pour évaluer combien bon est un code [BAT 1989a, BAT 1989b].

Etant donné un code en bloc linéaire (n, k, d) , on considère sa distribution de poids de Hamming comme un vecteur de probabilités $\underline{a} = (a_0, a_1, a_2, \dots, a_n)$ défini par $a_j := A_j / 2^n$ où A_j est le nombre de mots du code de poids j . Bien entendu, $a_j = 0$, $j < d$ et $a_d \neq 0$, et encore $\sum_{j=0}^n a_j = 1$.

D'après les résultats de Sidel'nikov [MacW&SLO 1977], la distribution de poids de Hamming d'un code binaire quelconque est telle que:

THEOREME (Sidel'nikov): Soit C un code binaire (n, k, d) ayant comme code dual C^\perp de distance de Hamming $d' \geq 3$. Alors $|A(z) - \phi(z)| \leq 20 / \sqrt{d'}$,

où $A(z) = \sum_{j \leq \mu - \sigma z}^n a_j$ (μ et σ sont la moyenne et l'écart type de la distribution),

et $\phi(z) = (2\pi)^{-1/2} \int_{-\infty}^z \exp(-t^2/2) dt = 1 - Q(z)$.

Ainsi, malgré leur distance minimale médiocre, la performance de quelques codes linéaires longs doit être assez proche de celle du code aléatoire. La difficulté consiste à trouver un décodeur souple dont la complexité soit raisonnable. Il conviendrait de remarquer que Cheung propose une relation explicite entre le caractère aléatoire d'un code q -aire et la distance minimale du code dual [CHE 1990]. Ainsi, qu'il s'agisse de codes binaires ou de codes q -aires ayant de fortes structures algébriques et de riches propriétés combinatoires, il est possible qu'il se comportent (en un certain sens) comme un code aléatoire, qui pourtant ne possède aucune structure.

La remise en question de la validité du critère de distance minimale revendiquée par Battail [BAT 1989] peut être généralisée à la modulation codée. Dans ce cadre, Battail et collaborateurs ont montré que la combinaison d'un code q -aire de Reed-Solomon RS (n,k) et d'une application des symboles de son alphabet sur une constellation plane symétrique de q points (par symétrique, il faut entendre que l'ensemble des distances entre l'un de ses points et tous les autres est indépendant du point considéré) a pour résultat une distribution des distances euclidiennes très voisine de celle que l'on obtient en moyenne par codage aléatoire, pourvu que q , n et k soient assez grands [BAT *et al.* 1990]. Plus tard, en utilisant l'approche du polynôme descripteur d'un code, Battail a été fondé à affirmer que le codage MDS constitue un moyen certain d'imiter le codage aléatoire [BAT 1991a].

On insiste sur le fait que la plupart des codes longs sont bons dans le sens qu'ils sont "proches" du code aléatoire et ont de ce fait une bonne performance. D'autre part, il y a très peu de codes ayant une forte structure mathématique qui sont bons au sens classique $d/n \rightarrow \epsilon > 0$ [COF&GOO 1990]. Cela explique bien la constatation désabusée "*All codes are good except those we can think of* - folk theorem". La démonstration de Shannon que la quasi totalité des codes longs sont bons ne fait pas intervenir le rapport d/n [SHAN 1949].

Il ne faut donc pas chercher de bons codes au sens classique car on butera sûrement sur la barrière "*except those we can think of*", très difficile à franchir. Pour l'ingénieur, les performances finales et la difficulté de mise en œuvre du décodeur sont les points décisifs du codage. Ces deux objectifs ne sont peut-être pas aussi antagonistes, voire incompatibles, qu'on l'a cru. Battail propose une direction pour sortir de cette impasse [BAT 1990b, BAT 1991a] en cherchant des codes fortement structurés qui imitent le codage aléatoire (il y en a probablement beaucoup - en tout cas, plus que les codes bons au sens classique). Ce processus comprend deux étapes: d'abord, une remise en question de la validité du critère de distance minimale. En effet, pour un code ayant un très grand nombre de mots, la performance n'est pas principalement déterminée par cette distance si le nombre des mots du code à la distance minimale d'un mot donné est négligeable. Ensuite, par la constatation que des codes bien structurés - donc probablement à "faible" distance minimale - peuvent imiter le codage aléatoire [BAT 89c, BAT *et al.* 1990, CHE 1989, CHE 1990]. La riche structure de ces codes peut conduire à des algorithmes de décodage (optimal ou sous-optimal) relativement simples, c'est-à-dire d'une complexité non prohibitive pour des longueurs importantes. Le nœud gordien pourra enfin être tranché!. Ces idées (bien que n'étant pas encore complètement divulguées ni acceptées dans une partie de la communauté scientifique) me semblent ouvrir de nouveaux horizons sur

l'avenir du codage. Néanmoins, le besoin de travailler avec des codes longs ne se fera sentir qu'à mesure qu'il existera une demande pour des débits au-delà du débit critique, ce que l'état de l'art actuel ne permet pas. Dans cette région difficile où le codage est impératif, les moyens traditionnels se sont avérés plutôt infructueux. L'avenir du codage de canal se jouerait-il à TELECOM PARIS ?. Cette thèse vient en quelque sorte appuyer et corroborer les idées ci-dessus.

1.2. APERÇU SUR LES CODES D'UNGERBOECK.

Les travaux d'Ungerboeck ont permis de "combiner modulation et codage" d'une façon logique, éclairante et nette. Le gain apporté par un tel procédé peut atteindre 3 à 6 dB en rapport porteuse à bruit en gardant la même bande.

La règle de décision d'un récepteur optimal (au sens de minimiser la probabilité d'erreur) sur un canal à bruit blanc gaussien additif consiste tout simplement à choisir, parmi l'ensemble de toutes les séquences de signaux permises, celle qui est la plus proche de la séquence reçue (bruitée) au sens de la distance euclidienne [VIT&OMU 1982].

Or, les erreurs les plus fréquentes correspondent, au moins pour les codes courts, à une décision erronée prise entre des séquences du code qui sont les plus proches. Autrement dit, une séquence émise étant donnée, il est plus probable que le bruit la change en l'une des séquences codées les plus proches qu'en une autre séquence ayant une distance euclidienne supérieure. Cette distance minimale s'exprime, dans un espace à nombre de dimensions infini engendré par un code convolutif, par:

$$d_{\min, \text{libre}}^2 = \min_{\{a_n\} \neq \{b_n\}} d_E^2(\{a_n\}, \{b_n\}) = \min \sum |a_n - b_n|^2, \quad \{a_n\}, \{b_n\} \in \mathbb{C}$$

où $\{a_n\}, \{b_n\}$ sont des séquences de longueur infinie appartenant au code.

Ungerboeck s'est donc proposé de maximiser cette distance minimale par un choix judicieux de l'espace de séquences permises (code). Le codage se fait par l'introduction d'interdépendance entre des séquences de symboles de la constellation constituante dans un espace à nombre de dimensions infini. Son idée consiste à augmenter la taille de la constellation non codée pour permettre d'ajouter la redondance et ensuite d'utiliser une partition en sous-ensembles du jeu de signaux élargi. Par exemple, pour une modulation MDP 4 non codée (2 bits/symbole), on considère une modulation MDP 8 codée, ayant donc 3 bits/symbole, dont 2 bits d'information et un troisième de redondance. Ainsi, le débit d'information est inchangé, de même que la bande passante requise.

Les sous-ensembles sont choisis de manière à augmenter la distance minimale à mesure que l'on progresse dans la partition. On a ainsi, par exemple, l'arborescence suivante:

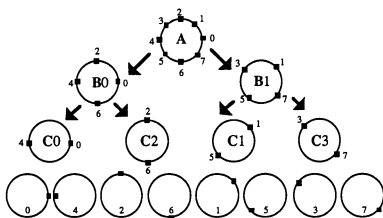


Figure 4. Illustration de la technique de partition d'Ungerboeck. Partition d'une constellation MDP 8.

On peut partager les huit signaux du niveau A en deux groupes de quatre éléments, notés B0 et B1 (plus généralement, la moitié des signaux dans chaque groupe). Ils correspondent aux constellations de deux modulations MDP 4 décalées l'une par rapport à l'autre. La distance minimale (euclidienne) que l'on obtient entre les signaux à l'intérieur d'un même groupe B (soit B0 ou B1) est

$$\Delta_1 = d_E(B) = \sqrt{E} \sin \pi/4 = \sqrt{2} \sqrt{E}.$$

Chaque groupe au niveau B de la partition peut à son tour être divisé en deux sous-groupes de deux signaux (encore la moitié des éléments), ce qui constitue le niveau C de la partition. La distance minimale (euclidienne) entre deux signaux quelconques à l'intérieur de chaque sous-groupe de C est donné par

$$\Delta_2 = d_E(C) = 2 \sqrt{E}.$$

En résumé, la distance minimale entre deux signaux à l'intérieur d'un même sous-ensemble à plusieurs niveaux de la partition (ici A,B et C) est donnée par:

$$\Delta_0 = d_E(A) = 2 \sqrt{E} \sin \pi/8 = 0,765 \sqrt{E},$$

$$\Delta_1 = d_E(B) = \sqrt{E} \sin \pi/4 = \sqrt{2E} \text{ et}$$

$$\Delta_2 = d_E(C) = 2 \sqrt{E}.$$

L'application bits/symbole MDP est déterminée par un triplet (y_0, y_1, y_2) tel que la première coordonnée spécifie le sous-ensemble B, la deuxième coordonnée détermine le sous-ensemble C à l'intérieur du sous-ensemble B préalablement sélectionné, et finalement la troisième coordonnée choisit un point de la constellation (symbole constituant à émettre) dans le sous-ensemble C déjà fixé par les coordonnées précédentes.

L'évolution du signal au cours du temps peut être représentée par un chemin dans un treillis, comme pour toute machine séquentielle. Dans son projet de codes en treillis, Ungerboeck impose quelques règles générales de symétrie (la symétrie apparaît encore comme un caractère essentiel des bons jeux de signaux):

primo- Tous les signaux apparaissent un même nombre de fois, avec symétrie (une même probabilité est attribuée à tous les signaux).

secondo- Les transitions dites parallèles sont associées à des signaux ayant entre eux une distance maximale (dans l'exemple, des signaux appartenant au même sous-ensemble C).

tertio- Les transitions associées aux branches arrivant à un même nœud du treillis, ou en partant, doivent correspondre à des signaux du même sous-ensemble (dans l'exemple, appartenant à un même sous-ensemble B).

Pour émettre m bits/symbole avec une modulation à deux dimensions, c'est-à-dire avec une efficacité spectrale de m (bits/s)/Hz, on utilise une constellation agrandie, à 2^{m+1} points, divisée en sous-ensembles (la modulation non codée ayant même débit utile n'exige que 2^m points). Ungerboeck a montré qu'un doublement de la taille de l'alphabet suffit pour atteindre la capacité [UNG 1982]. A présent, cependant, la modulation codée multidimensionnelle utilise une expansion moindre, comme on le verra plus en détail dans la suite.

Les diagrammes de treillis de quelques codes d'Ungerboeck sont représentés dans la figure ci-dessous.

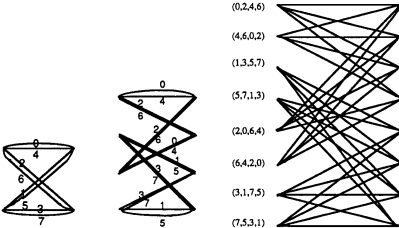


Figure 5. Diagramme de treillis à 2, 4 et 8 états pour les codes d'Ungerboeck MCT/MDP.

Dans une modulation non codée comme la MDP, le signal émis pendant un intervalle de temps n'a pas de rapport avec les signaux précédemment émis ni avec ceux qui le seront ensuite. Une modulation de ce genre est dite à états successifs indépendants, et toutes les combinaisons de symboles à deux dimensions sont permises. Dans la modulation codée, en revanche, le signal émis pendant un intervalle (associé aux éléments binaires émis pendant cet intervalle) dépend non seulement des éléments binaires d'entrée de cet intervalle, mais aussi du passé, c'est-à-dire de l'état du système. Ainsi, toutes les combinaisons de signaux ne sont pas permises et la règle selon laquelle on admet une séquence définit le code. On voit aisément que ce processus introduit de la redondance tout comme le codage classique. Par exemple, pour le code en treillis à deux états (fig. 5a), on voit que la séquence $s_0, s_6, s_1, s_2, s_7, s_7, \dots$ est une séquence du code alors que $s_0, s_1, s_1, s_2, s_7, s_6, \dots$ n'en est pas une.

Ici, on ne présente que des réalisations du codeur sous forme systématique avec rétroaction (fig 6) pour les treillis décrits dans la figure 5.

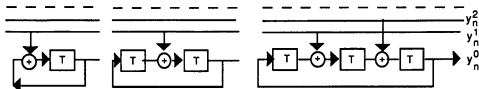


Figure 6. Codeurs d'Ungerboeck MCT/MDP 8 sous forme systématique avec rétroaction avec un treillis à 2,4 et 8 états, respectivement.

Les gains asymptotiques apportés par le codage, c'est-à-dire à des rapport signal à bruit très élevés, sont calculés en décibels par:

$$G_{\text{codé/non codé}} = 10 \log_{10} \left[\frac{d_{\text{libre, codé}}^2}{E_{s, \text{codé}}} / \frac{d_{\text{libre, non codé}}^2}{E_{s, \text{non codé}}} \right],$$

où d_{libre}^2 et E_s sont, respectivement, le carré de la distance euclidienne libre et l'énergie moyenne des signaux dans la constellation, affectés des indices "codé" et "non codé" pour identifier le système leur correspondant.

Ce gain est fréquemment interprété de la façon suivante:

$$G_{\text{codé/non codé}} = G - L,$$

où $G = 10 \log_{10} [d_{\text{libre, codé}}^2 / d_{\text{libre, non codé}}^2]$ est le gain en distance apporté par le codage et $L = 10 \log_{10} [E_{s, \text{codé}} / E_{s, \text{non codé}}]$ est la perte due à l'agrandissement de la constellation.

Le décodage à décision pondérée est effectué par l'algorithme de Viterbi avec un retard de décision de $M = 6v$ en prenant le carré de la distance minimale comme métrique.

Afin de donner une idée des gains et les complexités que l'on obtient en utilisant cette méthode, le tableau V présente les résultats de divers schémas MCT [UNG 1987].

TABLEAU V.

GAIN ET NOMBRE D'ETATS DU TREILLIS DE QUELQUES SCHEMAS DE MODULATION CODEE A DEUX DIMENSIONS.

| nombre d'états | MDP 8 /MDP 4 | MDP 16 /MDP 8 | MAQ 16 /MDP 8 | MAQ 32 /MAQ 16 | MAQ 64 /MAQ 32 |
|-------------------|-----------------|------------------|------------------|-------------------|-------------------|
| 4 | 3,01 | 3,54 | 4,36 | 3,01 | 2,80 |
| 8 | 3,60 | 4,01 | 5,33 | 3,98 | 3,77 |
| 16 | 4,13 | 4,44 | 6,12 | 4,77 | 4,56 |
| 32 | 4,59 | 5,13 | 6,12 | 4,77 | 4,56 |
| 64 | 4,59 | 5,33 | 6,79 | 5,44 | 5,23 |
| 128 | 5,17 | 5,33 | 7,37 | 6,02 | 5,81 |
| 256 | 5,75 | 5,51 | 7,37 | 6,02 | 5,81 |

Plus généralement, pour un système codé ayant une efficacité spectrale de m (bits/s)/Hz, on utilise une modulation à 2^{m+1} points combinée avec un codeur convolutif de taux $R = \tilde{m} / (\tilde{m} + 1)$, où $\tilde{m} \leq m$ dépend du nombre de niveaux de la partition que l'on considère. Les points de la constellation agrandie sont divisés en $2^{\tilde{m}+1}$ sous-ensembles, ayant chacun $2^{m-\tilde{m}}$ points:

$$2^{\tilde{m}+1} \text{ sous-ensembles} \cdot 2^{m-\tilde{m}} \text{ points / sous-ensemble} = 2^{m+1} \text{ points.}$$

Les meilleurs codes convolutifs de longueur de contrainte assez petite sont normalement construits en maximisant l'accroissement de distance de Hamming aux extrémités de l'événement d'erreur [CLA&CAI 19BB]. Autrement dit, la première et la dernière branches y contribuent majoritairement à la valeur totale de la distance entre un chemin correct et un chemin erroné.

Ungerboeck a cherché des schémas de MCT avec un accroissement important de la distance euclidienne entre les branches qui divergent puis convergent dans un événement d'erreur. D'une façon similaire, il a trouvé que les meilleurs codes en treillis sont obtenus en y maximisant le "poids" de la première et de la dernière branche.

Cependant, bien que ce comportement soit observé pour les codes à mémoire courte, il ne semble pas que le projet de codes en treillis MCT à longueur de contrainte élevée doive être guidé par le même critère. En cherchant à analyser les codes en treillis à très grand nombre d'états, il serait utile d'examiner la distribution des événements d'erreur (spectre de distances).

Il s'avère que la structure des événements d'erreur des codes longs n'est pas la même que celle des codes courts. Plus spécifiquement, la contribution à la distance libre minimale que l'on obtient aux extrémités d'un événement d'erreur peut ne pas être significative par rapport à la distance totale dans un événement d'erreur. Le projet de codes longs n'est pas une question complètement réglée (les 6 dB de gain obtenus par les codes en treillis correspondent à peu près au débit critique [FOR *et al.* 1984]). Cela ne sera probablement nécessaire qu'à des débits très élevés, dépassant même le débit critique du canal. Encore faudra-t-il changer le critère utilisé dans la construction des codes pour y réussir. Il semble que la distribution de distance va jouer un rôle prépondérant par rapport à la distance minimale. Là, encore, on rejoint les idées de Battail à propos de la construction explicite de bons codes longs.

I.3 ETAT DE L'ART.

I.3.1 GAIN POTENTIEL DE LA MODULATION CODEE.

La probabilité d'erreur par symbole dans une modulation MAQ non codée est bornée conformément à [PRO 1989]:

$$P_e(M) \leq 1 - [1 - \operatorname{erfc}(\sqrt{\frac{3\gamma_{av}}{2(M-1)}})]^2,$$

où $Q(x) = \frac{1}{2} \operatorname{erfc}(x/\sqrt{2})$ ou $\operatorname{erfc}(y) = 2Q(\sqrt{2}y)$.

Alors, pour des rapports signal à bruit assez élevés ($\gamma_{av} \gg M$), on a l'approximation suivante:

$$P_e(M) \approx 2 \operatorname{erfc}(\sqrt{\frac{3\gamma_{av}}{2(M-1)}}).$$

Tout d'abord, on remarque que l'agrandissement du jeu de signaux de façon à augmenter le débit d'un bit par symbole à deux dimension (1bit/symbole 2D) exige de doubler la taille de la constellation de départ, soit $M' = 2^{k+1}$ au lieu de $M = 2^k$. La probabilité d'erreur que l'on en tire est

$$P'_e(M') \approx 2 \operatorname{erfc}(\sqrt{\frac{3\gamma'_{av}}{2(M'-1)}}).$$

Ainsi, pour assurer le même taux d'erreur, la perte en puissance due à l'expansion de l'alphabet est d'environ 3 dB toutes les fois que l'on double la taille de la constellation. En effet, puisque $P'_e(M') = P_e(M)$, il s'ensuit que $\gamma'_{av} = 2\gamma_{av}$.

En cherchant à évaluer les gains potentiels de la modulation codée, on fixe d'abord un niveau souhaitable pour le taux d'erreur (disons 10^{-5}), et l'on examine ensuite quel rapport signal à bruit est nécessaire pour garantir ce taux.

Pour une modulation non codée il vaut, d'après les approximations ci-dessus,

$$\gamma_{av} \approx \frac{2[\operatorname{erfc}^{-1}(0,510^{-5})]^2}{3}(M-1).$$

Plus généralement, le rapport signal à bruit dont on a besoin pour maintenir le taux d'erreur inférieur à ϵ est donc

$$\gamma_{av} \approx \frac{2[\operatorname{erfc}^{-1}(0,5 \epsilon)]^2}{3} (M-1).$$

D'autre part, le rapport signal à bruit minimal pour atteindre ce taux d'erreur est, selon la formule de la capacité, $\gamma_{av} = (M-1)$, $M = 2^{2R}$ étant le nombre de points par deux dimensions. Forney *et al.* remarquent que le gain potentiel de la modulation codée, pour des taux d'erreur entre 10^{-5} et 10^{-6} , est de 8 (9 dB) [FOR *et al.* 1984].

Or, cette valeur a été mentionnée par Oliver, Pierce et Shannon il y a bien longtemps [OLI *et al.* 1949,p.1327], remarquant que la MIC exige à peu près une puissance supérieure d'un facteur 8 (9 dB) à la limite théorique établie par la capacité!

TABLEAU VI. GAIN POTENTIEL APPORTE PAR LA MODULATION CODEE.

| niveaux d'erreur ϵ | rapport signal-à-bruit (non codée) γ_{av} | gain (en dB) à $P_e = \epsilon$ |
|--------------------------------|---|------------------------------------|
| 10^{-4} | $5,48(M-1)$ | 7,39 |
| 10^{-5} | $6,95(M-1)$ | 8,42 |
| 10^{-6} | $8,42(M-1)$ | 9,25 |
| 10^{-7} | $9,91(M-1)$ | 9,96 |
| 10^{-8} | $11,40(M-1)$ | 10,57 |
| 10^{-9} | $12,90(M-1)$ | 11,11 |
| 10^{-10} | $14,40(M-1)$ | 11,58 |
| 10^{-11} | $15,90(M-1)$ | 12,01 |
| 10^{-12} | $17,41(M-1)$ | 12,41 |

Gain potentiel de la modulation codée

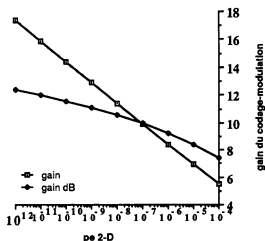


Figure 7. Gain potentiel du codage à divers niveaux de probabilité d'erreur.

On voit aisément que, bien que le gain de codage soit limité pour les taux d'erreur utilisés en pratique, le meilleur gain asymptotique, paramètre très employé pour la conception des codes, est en fait illimité.

La capacité d'un canal à bruit additif, blanc et gaussien peut être scindée en deux parties, l'une fonction seulement du rapport signal à bruit et l'autre ne dépendant que de la bande passante. Le codage n'est en fait lié qu'à la première partie [WOZ&JAC 1967, GAL 1968]: $C_N = 1/2 \log_2(1 + S/N)$ bits/dim. D'autre part, la limitation de bande impose aussi une limitation du débit [WYN 65, JER 77] exprimée par $C_B = 2B$ dim/s. La capacité du canal gaussien est donc donnée par $C = C_B C_N = B \log_2(1 + S/N)$ bits/s.

Pour aborder le codage, il semble raisonnable d'exprimer la capacité normalisée à deux dimensions: $C_N^* = \log_2(1 + S/N)$ bits / symbole 2D ou (bits/s)/Hz. D'ailleurs, on interprète assez souvent la limitation de débit due à la bande comme étant $C_B^* = B$ symboles 2D/s = B bauds (la capacité peut aussi bien être exprimée en (bits/s)/bauds).

Il y a deux points importants qui doivent être éclaircis. D'abord, les techniques de modulation codée ne conviennent pas qu'aux canaux à bande étroite, mais toutes les fois

que le canal appartient à la région limitée en bande. Ainsi, elles peuvent être fort utiles pour des canaux typiquement à bande large tels que les canaux de satellite. En effet, même si un accroissement de la puissance disponible dans les répéteurs a été réalisé, les exigences de débit ont augmenté encore bien plus vite, de sorte que la limitation de bande s'y fait sentir. Ensuite, on doit prendre garde à la proximité entre le taux d'émission déjà atteint dans la pratique et la capacité. Les systèmes pratiques mis en œuvre jusqu'ici n'ont pas encore atteint des débits aussi proches de la capacité que l'on aurait pu le penser (voir l'exemple 10.3.1 dans le récent ouvrage "Elements of Info. Theory", p.250, T. COVER et J. THOMAS, Wiley, 1991). La comparaison doit toujours être faite en tenant compte de la bande de fréquences utilisée. Par exemple, la valeur de 19 200 bits/s ne correspond au débit critique du canal téléphonique que si la largeur de bande utilisée est de 2 400 Hz. La limite ultime (le débit critique) proposée en [FOR *et al.* 1984] s'écrit plus clairement comme une efficacité spectrale de 8 (bits/s)/Hz. Si l'on utilise une bande plus large (comme dans les MODEM d'aujourd'hui → 2 743,86 Hz), un débit de 19 200 bits/s est équivalent à une efficacité spectrale de 7 (bits/s)/Hz: il ne correspond pas exactement au débit critique du canal. Cependant, cette réussite représente sans l'ombre d'un doute un formidable développement des débits numériques. Dans les années soixante-dix, personne ne croyait qu'il serait possible de parvenir à tels débits sur les voies téléphoniques [HOL 1977]. Pour éviter les difficultés dans l'évaluation de la proximité du débit et de la capacité pour différentes largeurs de bande, il convient de l'exprimer en fonction de l'efficacité spectrale, comme on le montre dans la figure suivante.

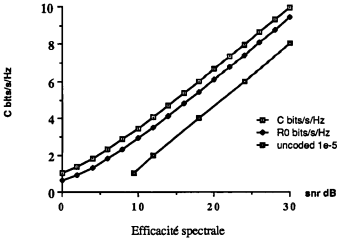


Figure 8. Capacité et débit critique exprimés en termes de limitation de l'efficacité spectrale.

La transmission de données sans interférence entre symboles est possible même dans des canaux idéalement limités en bande utilisant des filtres de Nyquist, pourvu que le débit de Nyquist ne soit pas dépassé [BIC *et al.* 1986, PRO 1989, etc].

La largeur de bande impose donc une borne supérieure au débit et cette limitation ne découle pas de la théorie de l'information.

Bien que l'on en déduise la limite $2B$ dim/s, il faut mentionner quelques autres (étonnants) résultats récents. Ils montrent qu'il est possible d'émettre à un débit 25% plus grand que le débit de Nyquist sans réduire la distance minimale, malgré la présence d'interférence entre symboles [MAZ 1975, MAZ&LAN 1988, HAJ 1990]. Ces résultats concernent exclusivement des problèmes liés à la limitation de bande et n'ont aucun lien avec le rapport signal à bruit. Quoique ces résultats n'aient pas d'incidence sur le codage (ni la capacité par dimension de Shannon), il peuvent s'avérer profitables pour améliorer le débit d'information au-delà de la capacité en bits/s (pas en bits/dimension!). C'est certainement une voie de recherche qui doit être envisagée dans l'avenir.

I.3.2 MODULATION CODEE EN TREILLIS (MCT).

L'utilisation du codage correcteur d'erreur direct (FEC) est maintenant très répandue et largement acceptée pour augmenter l'efficacité de la transmission dans virtuellement dans tous les canaux de communication numérique. La situation est un peu plus compliquée dans la MCT car l'application non linéaire bit/symbole rend le codeur partiellement non linéaire quoique le code convolutif employé soit linéaire.

Thapar a examiné les performances des codes d'Ungerboeck dans un canal gaussien en présence de diverses distorsions y compris les distorsions linéaires, non linéaires, gigue de phase, écart de fréquence et bruit additif [THA 1984]. Il a montré une fois pour toutes, au moyen d'une mise en œuvre pratique sur un processeur de signaux numérique à 12 bits, que les codes d'Ungerboeck sont très robustes et peuvent réaliser une grande amélioration des performances même en présence de distorsions sévères.

En 1986, Wilson et d'autres ont présenté la conception et le projet d'un codage hybride entre modulation codée et codage classique dans le sens que le codage apporte un élargissement partiel de bande [WIL *et al.* 1986]. Ils ont conçu un codage en treillis basé sur un code convolutif de taux 5/6 à 8 états dont le gain sur un canal à bruit additif gaussien est de 6 dB par rapport au système MDP 8 au prix d'une augmentation de 16% de la bande.

D'autres aspects théoriques assez intéressants sur le codage-modulation ont été abordés par Benedetto, Marsan et collaborateurs [MAR *et al.* 1984, voir aussi BEN *et al.* 1988]. Des codes linéaires appartenant à une classe particulière nommée codes superlinéaires ont été introduits. Leur propriétés permettent d'épargner largement le temps de calcul lors de l'analyse des performances. Une recherche de codes MCT sur les MDP 16 et MDP 32 dont la complexité va de 4 à 128 états a été réalisée et a abouti à de nouveaux codes.

Wilson et collaborateurs ont étudié un système de modulation codée pour le canal de satellite ayant une efficacité spectrale de 3 (bits/s)/Hz [WIL *et al.* 1984]. Le codage de canal consiste en la combinaison d'un code convolutif à 32 états avec la modulation MDP 16; le gain apporté est d'environ 4,8 dB. La performance de ce système est évaluée sur un canal non linéaire fortement limité en bande. Bien que la performance du système codé soit dégradée par rapport au système non codé, on montre que le gain du codage est grosso modo préservé.

Benedetto *et al* examinent le projet de paramètres critiques du récepteur pour un système de codage-modulation MCT / 16 MDP dans l'intention de limiter la complexité au minimum possible en vue d'une construction en VLSI. Le retard de décodage de l'algorithme de Viterbi, l'effet de la quantification des métriques et la présence de gigue de phase dans la porteuse récupérée ont été évalués par simulation [BEN *et al.* 1989].

Viterbi, Wolf et collaborateurs proposent un très intéressant schéma pour l'implémentation commerciale des codes en treillis [VIT *et al.* 1989]. Le principal avantage de cette approche est de permettre la mise en œuvre d'une grande variété de codes MCT quasi-optimaux à l'aide d'un simple et unique codeur-décodeur de taux 1/2, tout simplement en opérant des changements mineurs de connexion pour introduire des branches parallèles. Il ont obtenu ainsi des gains allant de 3 dB à environ 6 dB, en fonction du codeur sélectionné.

Le décodage souple des codes MCT peut être fait soit au moyen de l'algorithme de Viterbi (optimal) [VIT&OMU 1982], soit par un algorithme séquentiel (versions de Fano-Wozencraft ou aidé par un génie, etc) [AND&MOH 1984], voir aussi [POT&TAY 1989].

L'application de la modulation codée aux transmissions par satellite, plus spécifiquement des codes MCT sur la modulation MDP, a été réexaminée de façon remarquable par Hagenauer [HAG 1989]. Cela englobe d'une part la M-MDP codée employée dans le système INTELSAT et d'autre part celle qui est appliquée aux services mobiles du système M-SAT. Des canaux INTELSAT à 64 kbits/s pour la transmission de voix et de données utilisent un code d'Ungerboeck à 4 états sur la MDP 8 tandis que des canaux à 120 Mbits/s (140 Mbits/s) possèdent actuellement un modem 8 MDP codée en MCT à 16 états. Il est à remarquer que la synchronisation et la récupération de phase de la porteuse consomment près de la moitié de la puissance des processeurs.

Il est connu que, pour des rapports signal à bruit modérés, les performances des codes peuvent être prédites à partir du spectre des distances et pas simplement de la distance minimale (même compte tenu du nombre des mots à cette distance). Rouanne a présenté un algorithme tiré de l'algorithme bidirectionnel à pile destiné à l'évaluation du spectre de distance euclidienne des codes MCT [ROU 1989].

Des gains assez modestes par rapport à la MCT classique peuvent être obtenus en introduisant une modulation asymétrique [DIV *et al.* 1987]. Cependant, la complexité qui en résulte ne justifie guère l'emploi de cette méthode.

Pottie et Taylor ont montré, en faisant appel au concept de poids euclidien d'Ungerboeck, comment déterminer exactement la distance minimale libre des codes MCT sur des modulations MAQ carrées en ne prenant en considération que les distances des chemins par rapport au chemin tout zéro [POT&TAY 1987].

Rouanne définit la classe des codes quasi-réguliers [ROU 1988, ROU 1990] comme l'ensemble des codes MCT dont le spectre de distance euclidienne peut être calculé à partir du nombre de chemins (mots de code) par rapport au chemin nul du treillis.

L'application de la modulation codée n'est pas restreinte aux modulations traditionnelles MDP/MAQ. Padovani et Wolf ont considéré le projet de MCT utilisant la technique de partition d'Ungerboeck avec une modulation numérique combinée en fréquence et en phase [PAD&WOL 1986]. Cette approche a l'avantage de limiter le nombre d'états de la constellation constituante. Par exemple, il est possible de concevoir une modulation 4-MDF / 8-MDP codée pour obtenir des gains de codage importants avec une efficacité spectrale de 4 (bits/s)/Hz (4 bits/T). On considère aussi le codage en treillis MCT appliqué à une modulation combinée 4-MDF/16-MAQ ayant une efficacité de 5 (bits/s)/Hz. Les caractéristiques spectrales de tels schémas sont évaluées et l'on montre qu'ils sont bien adaptés aux canaux fortement limités en bande passante.

Une borne inférieure de la meilleure distance minimale (libre) euclidienne des codes MCT a été développée par Rouanne [ROU 1987, ROU 1990]. Cette borne est déduite du codage aléatoire appliqué à l'ensemble des codes en treillis non linéaires. Il démontre que la distance libre augmente linéairement avec la longueur de contrainte pour des valeurs élevées de cette dernière. Etant une borne inférieure, elle garantit l'existence de bons codes MCT.

Calderbank et autres ont étudié les bornes de la distance euclidienne minimale qui peut être atteinte avec la modulation en treillis. Etant donné le nombre v de registres de mémoire, la distance minimale normalisée d'un code un treillis quelconque (v,k) ayant ses valeurs sur l'espace euclidien R^n obéit à la borne suivante [CAL *et al.* 1985]:

$$\frac{d_{\min}^2}{P(C)} \leq \begin{cases} \frac{3.285v}{4k/n} & \text{si } k/n \geq \ln^{-1}4 \\ \frac{0.870v}{k/n} & \text{si } k/n < \ln^{-1}4, \end{cases}$$

où $P(C)$ est la puissance normalisée à deux dimensions.

L'évaluation des performances d'une certaine classe de MCT est abordée par Wolf et Zehavi en exhibant une méthode de calcul de la distribution de poids des erreurs utilisant la fonction génératrice [ZEH&WOL 1988].

Honig suggère une technique pour réduire le temps de recherche (par ordinateur) de codes en treillis MCT de longueur de contrainte élevée en utilisant des connexions de treillis en parallèle [HON 1988].

Divsalar et Simon ont, de leur côté, inventé une nouvelle technique plus générale que la MCT, appelée Modulation Multiple Codée en Treillis (MMCT) où, contrairement aux systèmes MCT conventionnels, plusieurs symboles du canal sont attribués à chaque branche du treillis [DIV&SIM 1987, DIV&SIM 1988a]. Cette idée permet de parvenir très aisément à des efficacités spectrales fractionnaires. De plus, elle apparaît comme un outil intéressant dans le cas des canaux à évanouissement [DIV&SIM 1988c].

D'autres aspects particuliers du projet de MCT:

Bien des travaux ont été déjà consacrés au problème de l'emploi de la MCT en présence d'évanouissement. Un des plus sérieux inconvénients à cette application est lié au besoin d'une détection cohérente pour atteindre leur gain potentiel.

Makrakis et Feher proposent un algorithme séquentiel pour la détection différentielle de systèmes de MCT sur une modulation MDP [MAK&FEH 1988]. Le problème de la détection différentielle des codes MCT a été abordé pour la première fois par Simon et Divsalar [SIM&DIV 1987], puis largement considéré par Edbauer [EDB 1987, EDB 1989a, EDB 1989b]. Il a évalué les performances d'un codage-modulation avec entrelacement sur une modulation MDP différentielle (8-DPSK) dans un canal à évanouissement de Rayleigh ou de Rice. McLane et collaborateurs ont étudié l'application et les performances de la MCT pour la transmission de la voix sur un canal mobile par satellite en considérant la MDP et la MDP différentielle codée en treillis [McL *et al.* 1987, McL *et al.* 1988]. Lee et Farrel ont évalué par simulation les performances de codes en treillis MDP/MCT à 4 et 8 états avec invariance de phase, et dont la longueur de contrainte est 3 et 4, sur un canal avec évanouissement de Rayleigh [LEE&FAR 1991]. Au taux d'erreur de 10^{-3} , les gains obtenus sur la modulation non codée correspondante sont de 2,5, 3,7, 4,0 et 5,1 dB. Moher et Lodge proposent l'utilisation d'une porteuse pilote avec les données codées pour permettre la détection cohérente sur un canal de Rice avec évanouissement [MOH&LOD 1989]. Cette méthode s'avère très robuste sur une vaste gamme de canaux à évanouissement.

Du coté théorique, Divsalar et Simon ont déduit une borne de Chernoff (exponentiellement fine) sur la probabilité d'erreur des signaux deux à deux pour des codes MCT sur un canal gaussien avec évanouissement de Rayleigh [DIV&SIM 1988b]. A partir de cette borne, il est possible de proposer un critère pour la construction de codes MCT adaptés aux canaux à évanouissement, en utilisant encore l'idée d'une partition des ensembles [DIV&SIM 1988c]. Des résultats similaires ont été trouvés par Schlegel et Costello [SCH&COS 1989]. Divers nouveaux codes mieux appropriés à ce genre de canaux ont été trouvés [SCH&COS 1989, JAM&LE_N 1987]. Néanmoins, dans le cas des communications de radio mobile, la fonction d'autocorrelation du fading est une fonction de Bessel dont la décroissance est lente. Il est alors difficile de réaliser un entrelacement parfait pour rendre les échantillons non corrélés, à cause de la très grande durée du retard. Le problème de la conception et des performance de la MCT sur un canal à évanouissement de Rayleigh avec mémoire a été généralisé par Leeuw-Boulé et autres [LEEU *et al.* 1991].

En général, on n'a considéré pour cette application que des codes MCT sur une modulation MDP, compte tenu de leur propriété d'enveloppe constante. Du et Vucetic mentionnent qu'avec l'utilisation d'une porteuse pilote, les modulations codées sur MAQ peuvent présenter des avantages sur la MDP codée [DU&VUC 1991]. Ils proposent donc des codes MCT sur la modulation MAQ 16 pour le canal à évanouissement de Rayleigh. Le développement du codage MCT pour la transmission mobile par satellite est considéré dans [YUE *et al.* 1990].

Les performances et le projet de codes MCT sur des canaux non linéaires à mémoire finie ont été abordés en utilisant les séries orthogonales de Volterra pour modéliser un système non linéaire discret. [LIU *et al.* 1990]. Des schémas MCT sur le canal de satellite avec des non-linéarités ont été étudiés par Biglieri, et Marsan et collaborateurs [BIG 1984, MAR *et al.* 1987]. Le codage sans élargissement de bande sur la modulation 16-MDP pour transmettre 3 (bits/s)/Hz donne un gain de 2 dB au taux d'erreur de 10^{-6} en considérant les effets de la limitation de bande et des non-linéarités.

Il est parfois utile d'admettre la présence d'une interférence entre symboles contrôlée en vue de la mise en forme spectrale du signal, par exemple spécifiant la position de zéros de la réponse de fréquences [KAB&PAS 1975]. Cela est utile dans le cas d'une forte limitation de bande. D'autre part, les modèles de canaux à réponse partielle sont aussi très utiles pour l'enregistrement numérique magnétique. Les besoins du codage correcteur d'erreur se sont beaucoup accrus avec l'avènement d'enregistrements à très haute densité visant à répondre à la demande toujours croissante dans les domaines du stockage de données. Le projet de codes du type MCT adaptés à des canaux à réponse

partielle est abordé dans les références [WOL&UNG 1986, ZEH&WOL 1988 et FOR&CAL 1989].

Le projet de codes MCT pour la transmission numérique en présence d'interférence entre symboles est aussi considéré dans [EYU&QUR 1989, WON&McL 1988].

I.3.3. MODULATION CODEE EN BLOC (MCB).

Beaucoup d'intérêt a été porté, dès la dernière décennie, aux codes où la redondance n'est pas obtenue en ajoutant des symboles de parité, mais par l'intermédiaire d'une augmentation de la taille de l'alphabet. Le travail d'Ungerboeck étant fondé sur des codes convolutifs, on a été naturellement conduit à se poser la question d'adapter ces techniques aux codes en bloc.

Un des pionniers de la modulation codée en bloc (MCB) est Cusack qui a introduit une application de la technique de partition d'Ungerboeck aux codes de Reed-Muller combinés à une modulation MAQ [CUS 1984].

Forney, Gallager et collaborateurs ont bien explicité les moyens de réaliser la partition des ensembles pour appliquer la modulation codée à des codes en bloc.[FOR *et al.* 1984]. Ils ont présenté la mise en œuvre d'une famille de schémas basés sur des codes en bloc connus, pour les dimensions $N = 4, 8, 16$ et 24 , conduisant à des gains asymptotiques de $1,5, 3,0, 4,5$ et $6,0$ dB, respectivement. Le processus de décodage souple (à maximum de vraisemblance) peut être réalisé à l'aide d'un treillis. Dans ce remarquable article, le rapport étroit existant entre le codage-modulation en bloc et la construction de réseaux denses est clairement mis en évidence. C'est principalement cette dernière (riche) interprétation du codage-modulation en bloc qui est adoptée dans cette thèse.

Ensuite, Sayegh a généralisé l'approche du codage MCB proposée par Cusack et a définitivement établi la puissance des codes en bloc dans le domaine de la modulation codée [SAY 1986]. Cette généralisation a consisté à appliquer divers codes binaires connus au projet de systèmes de modulation-codage adaptés à plusieurs constellations de signaux MDP et MAQ.

On illustre dans ce qui suit l'idée de modulation codée en bloc par des exemples très simples du projet conjoint d'un codage en bloc avec une constellation constituante à 2 dimensions ayant 2^m points comme dans la figure ci-dessous.

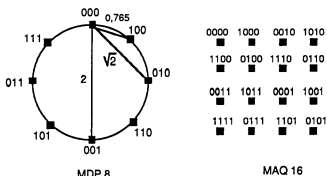


Figure 9. Constellations MDP et MAQ et leurs applications respectives bits/symbole.

On forme une matrice d'éléments binaires ayant n colonnes et m lignes. Chaque colonne correspond à un point de la constellation à 2 dimensions selon l'application indiquée. L'ensemble de n points de l'espace à deux dimensions constitue un signal dans un espace à $2n$ dimensions (où chaque symbole est représenté par un point). Si l'on introduit de la redondance dans chaque ligne de la matrice, alors les n symboles à deux dimension qui constituent le signal multidimensionnel sont codés, c'est-à-dire que toutes les séquences de n symboles bidimensionnels ne sont pas permises. On prend donc autant de codes binaires que de lignes de la matrice, soit m codes constituants, tous de même longueur n . Une matrice binaire $n \times m$ correspond à un mot de code dans l'espace multidimensionnel si et seulement si chaque ligne est un mot du code binaire correspondant à cette ligne. Par exemple, pour la modulation MDP 8, on a $m = 3$. Soit (n, k_i, d_i) le code binaire qui détermine la ligne i , $i = 1, 2, 3$. Bien entendu, $0 < k_i \leq n$. En supposant que l'on impose la contrainte $k_1 + k_2 + k_3 = 3rn$, le débit d'information (utile) est alors $3r$ bits par symbole à deux dimension (toujours inférieur à 3 bits/symbole puisqu'il y a de la redondance). Pour r fixé, le problème consiste à choisir les codes $(k_i, i = 1, 2, 3)$ de façon à maximiser la distance euclidienne minimale qui en résulte. Cela est réalisé en tenant compte des propriétés suivantes de la distance euclidienne:

- Si deux mots du code diffèrent dans la première ligne, alors sûrement la distance entre eux est $\Delta \geq 0,765$ quelles que soient les lignes suivantes.

Si deux mots du code coïncident dans la première ligne, alors sûrement la distance entre eux est $\Delta \geq 1,414$ quelle que soit la dernière ligne.

Finalement, si les deux mots de code des deux premières lignes coïncident, alors $\Delta \geq 2$.

Le carré de la distance euclidienne minimale entre mots de code dans l'espace de signal est de ce fait

$$d^2_{\min} = \text{MIN} (0,586 d_1, 2d_2, 4d_3).$$

La modulation codée la plus efficace est déterminée par une recherche exhaustive des codes binaires (n, k_i, d_i) solutions du problème d'optimisation:

$$\begin{array}{ll} \text{MAX} & \text{Min} (0,586 d_1, 2d_2, 4d_3), \\ (n, k_i, d_i) & \\ \text{contraintes} & \\ 0 < k_i \leq n, & \\ k_1 + k_2 + k_3 = 3rn. & \end{array}$$

à l'aide des tableaux de bornes sur la distance de Hamming minimale pour des codes en bloc binaires [p. ex., HEL&STI 1973, VER 1987].

En 1988, Isaksson a développé une classes de codes MCB basée sur une combinaison de codes cycliques non-idéaux (définis par des matrices circulantes) avec la modulation MDP [ISA 1988]. Le décodage souple est effectué par l'algorithme de Viterbi à l'aide du treillis [WOL 1978]. Des résultats de simulation ont été présentés pour divers codes de taille moyenne; ils montrent que le gain de codage est, à des rapports signal à bruit modérés, très sensible aux nombre des mots du code à la distance minimale [ISA 1988, ISA 1989].

L'évaluation des performances de divers schémas MCB a été analysée par Kasahara et Du [KAS&DU 1988]. Huber a examiné un système de modulation codée en bloc à synchronisation facile basé sur l'application de codes de Reed-Solomon [HUB 1989].

Chouly et Sari ont récemment développé des familles de MCB basées sur le premier et le deuxième niveau de la partition dont le gain asymptotique est de 3,0 dB et 4,5 dB respectivement [CHO&SAR 1990]. Ces schémas, très attrayants pour des applications à haut débit, sont obtenus en combinant des codes en bloc de longueur très courte à la modulation MAQ. Des constellations MAQ à 32, 128 et 512 états ont été considérées, permettant d'obtenir des efficacités spectrales de 4, 6 et 8 (bits/s)/Hz.

Testard a présenté deux codes MCB adaptés à la transmission par faisceaux hertziens à très grand débit [TES 1990]. Dans ce cas, on cherche à combiner la modulation MAQ 128 avec des codes simples. L'intérêt de cette technique réside dans la simplicité des décodeurs. En effet, la complexité du décodeur de Viterbi dans les récepteur MCT limite

leur application à des systèmes à débits moyens, inférieurs à 45 Mbits/s. Les résultats sont comparés à ceux d'une modulation MAQ 64 non codée. Plusieurs procédés de décodage connus [FOR *et al.* 1984, SAY 1986, etc.] sont abordés et le compromis entre complexité et gain est évalué. Le même problème a été récemment abordé par Karam et collaborateurs en considérant le codage en bloc multiniveau [KAR *et al.* 1991]. Un schéma de modulation codée pour la MAQ 128 basé sur le deuxième niveau de partition est présenté; son gain asymptotique est d'environ 4 dB à 10^{-8} et la complexité du décodeur assez modérée.

De son côté, Chang a appliqué la technique de partition d'Ungerboeck au projet d'une modulation codée en bloc dans le but d'augmenter l'ordre de diversité de l'émission sur un canal à évanouissement de Rayleigh non sélectif [CHA 1988]. Il a comparé la performance d'un système MDP 8 codée en bloc avec celle d'un système de radio mobile conventionnel utilisant un code convolutif binaire de taux 2/3 appliqué à une modulation MDP 4. On constate qu'une diversité de même ordre est possible sans élargissement de la bande.

L'idée du codage en plusieurs étapes, conçue en 1977 par Imai et Hirakawa, consiste en une combinaison de plusieurs codes correcteurs d'erreur selon une partition de la constellation des signaux en sous-ensembles [IMA&HIR 1977]. Elle est fortement liée à la modulation codée.

Portugheis considère la combinaison de la modulation MDP avec des codes en bloc et exhibe une famille de codes linéaires et invariants à la rotation de phase dont le décodage peut être réalisé par un décodeur à plusieurs étages [POR 1991].

La construction de MCB adaptées aux canaux à évanouissement de Rayleigh a été aussi étudiée. Vucetic et collaborateurs considèrent la modulation MDP-M combinée à des codes en bloc sur un anneau modulo M [VUC *et al.* 1990]. Il y est montré que le critère pour le projet de tels codes doit tenir compte non seulement de la distance minimale euclidienne mais aussi de la longueur effective des chemins du treillis [aussi ZHA&VUC 1990].

La conception de systèmes efficaces de modulation codée en bloc faisant appel à la théorie des codes de classes latérales de Forney [FOR 1988] a été abordée dans [KSC *et al.* 1989]. Ils ont ainsi abouti à une interprétation du codage MCB combiné à la modulation MDP qui est entièrement basée sur des propriétés des chaînes de partition.

I.4. THEORIE MODERNE DE LA MODULATION CODEE: QUELQUES SUJETS AVANCES.

I.4.1 MODULATION CODEE EN TREILLIS MULTIDIMENSIONNELLE.

Le codage MCT multidimensionnel est une nouvelle méthode de codage qui permet de réaliser des gains significatifs sans trop augmenter le nombre de signaux dans la constellation agrandie. Le premier code MCT multidimensionnel, construit sur une modulation à 4 dimensions, a été proposé par Gallager [FOR *et al.* 1984]. En 1985, Calderbank et Sloane développaient indépendamment une famille similaire [CAL&SLO 1985]. Cependant, ce sujet n'a suscité beaucoup d'attention qu'après la publication de l'article de Wei où étaient présentés divers schémas de MCT à dimensions 4, 8 et 16 [WEI 1987]. En outre, Ungerboeck dans son article à l'IEEE Comm. Magazine envisage aussi la modulation codée multidimensionnelle [UNG 1987b]. Forney a classé les codes MCT multidimensionnels selon une classification géométrique en classes latérales [FOR 1988]. Wei a repris le problème de la recherche de codes invariants à l'ambiguïté de phase et a découvert une famille de codes MCT/MDP multidimensionnels [WEI 1989]. La modulation codée en treillis multidimensionnelle sur une modulation de phase a aussi été récemment abordée par Pietrobon, Ungerboeck et autres [PIE *et al.* 1990].

Dans la MCT classique (c'est-à-dire un codage en treillis sur une modulation à deux dimensions), un bit de redondance est ajouté à chaque symbole non codé, ce qui a pour résultat une constellation agrandie dont la taille est double de celle de départ. Or, le coût en puissance moyenne exigé pour cette opération est d'environ 3 dB (se reporter à I.3.1). Cette perte doit donc être compensée par le gain du codage. Dans la MCT multidimensionnelle, on ajoute un bit de redondance toutes les $2n$ dimensions, autrement dit, à chaque bloc de n symboles MAQ. Par conséquent, on réduit le nombre de bits de redondance par symbole bidimensionnel et donc l'augmentation de la taille de la constellation constituante dont on a besoin pour le système codé. Pour une modulation codée MCT basée sur une modulation à 4 dimensions, le prix à payer dû à l'expansion de l'alphabet est réduit à 1,5 dB, tandis que pour une MCT à 8 dimensions cette perte devient 0,75 dB environ. De plus, cette réduction de taille de la constellation constituante à deux dimensions permet de parvenir à une meilleure performance en présence d'autres distorsions que le bruit additif, telles que des non-linéarités ou giges de phase. Cependant, il doit être clair, quand on adopte cette stratégie, que le gain en distance euclidienne apporté par le codage (compte non tenu de la perte due à l'agrandissement des

constellations) est d'autant plus petit que la redondance introduite a été plus petite. Le gain "net" du codage, étant le résultat du bilan entre ce gain en distance et la perte due à l'expansion, peut être supérieur. Il est évident qu'il existe un compromis et que le gain final résultant du codage n'est pas d'autant plus grand que la dimension de la modulation constituante est grande.

Il convient de rappeler que le codeur convolutif ne travaille pas à la cadence de modulation dans le cas d'une modulation codée en treillis multidimensionnelle. Pour une MCT sur modulation à $2n$ dimensions, le codeur ne travaille qu'à un rythme n fois inférieur à la cadence de modulation.

Calderbank et Sloane ont présenté deux familles de MCT multidimensionnelles à huit états conçues sur une modulation à quatre dimensions dont les gains asymptotiques sont d'environ 5 dB [CAL&SLO 1985]. La première famille permet la transmission de 8 bits par symbole à quatre dimension, soit une efficacité spectrale de 4 (bits/s)/Hz, avec un gain de 4,7 dB sur la transmission non codée MAQ 16. Elle est proposée comme un moyen de codage adapté aux modems à 9 600 bits/s sur le canal téléphonique à bande 2400Hz. Quant à la deuxième famille, elle émet 12 bits par symbole à quatre dimensions et peut atteindre une efficacité spectrale de 6 (bits/s)/Hz; elle est ainsi appropriée aux modems à 14 400 bps sur le canal téléphonique.

Wei a analysé plusieurs codes MCT multidimensionnels sur des constellations constituantes à 4, 8 et 16 dimensions [WEI 1987]. Il a montré comment sélectionner les constellations et comment réaliser une partition des ensembles telle que Ungerboeck l'a faite pour la MCT sur une modulation à deux dimensions. Le codage MCT sur une modulation multidimensionnelle permet de réduire la taille de la constellation bidimensionnelle utilisée et partant d'obtenir une meilleure tolérance aux ambiguïtés de phase. D'ailleurs, ces codes présentent aussi un meilleur compromis entre la complexité et le gain de codage.

Tretter a aussi introduit une MCT multidimensionnelle invariante à une rotation de 90° de la phase, basée sur un code en treillis à 64 états qui permet la transmission de 4 bits/baud (4 (bits/s)/Hz) avec une constellation constituante d'à peine 20 points au lieu de la MAQ 16 [TRE 1989]. Ce codage atteint un gain de 5,23 dB par rapport à la modulation non codée 4X4 MAQ et même 1,23 dB par rapport au code en treillis de la norme V.32 du CCITT [CCITT 1989].

I.4.2. CODES DE RESEAUX ET DE CLASSES LATERALES

"Trellis codes are to lattice codes as convolutional are to block", G.D. FORNEY.

I.4.2.1 RESEAUX

"Lattices can be regarded as generalization of block codes, using Euclidean rather than Hamming weights". G.D. FORNEY.

Soit B_n une sphère à n dimensions centrée sur l'origine, de rayon ρ et de contenu $V(B_n) = V_n \rho^n$, où $V_n := \pi^{n/2} / \Gamma(\frac{n}{2} + 1)$ est le volume d'une hypersphère de rayon unité [p.ex. WOZ&JAC 1967].

Définition (*empilements*).

Soit $\{\underline{v}_1, \underline{v}_2, \dots\}$ une suite infinie de vecteurs dans l'espace euclidien R^n . On dit que l'ensemble $\{\underline{v}_1 + B_n, \underline{v}_2 + B_n, \dots\}$ des sphères décalées par ces vecteurs forme un empilement si et seulement si elles sont disjointes deux à deux. On s'intéresse en particulier au cas où la sphère B_n a le plus grand rayon possible compatible avec la condition que les sphères décalées sont disjointes, ce qui implique un contact entre les sphères voisines.

Un empilement est donc un arrangement d'une infinité d'hypersphères disjointes de même taille (voir [SLO 1984]). Forney mentionne à juste titre [FOR 1990] qu'il était clair, dès les travaux originaux de Shannon, que le problème de parvenir aux empilements les plus denses dans des espaces à grand nombre de dimensions est essentiellement le même que d'atteindre la capacité dans un canal à bruit blanc gaussien.

La densité Δ d'un empilement est la fraction de l'espace couverte par les sphères (un terme mieux approprié pour désigner cette grandeur pourrait être "coefficient de remplissage"). Il va sans dire que $\Delta \leq 1$. Le problème de l'empilement de sphères consiste alors à choisir la position de leurs centres de façon à maximiser la densité qui en résulte.

Une sphère quelconque de centre \underline{x} appartenant à l'empilement est en contact avec un certain nombre d'autres sphères, dont le nombre est appelé le nombre de contacts (*kissing number*) du réseau, noté $\tau(\underline{x})$.

La densité de centres, δ , est la densité divisée par le volume d'une sphère de rayon unité, $\delta := \Delta / V_n$, c'est-à-dire le nombre des centres de sphères de l'empilement par unité de volume.

On simplifie grandement la recherche des empilements denses de sphères dans les espaces multidimensionnels en ne considérant qu'une classe particulière d'empilements qui ont une configuration très régulière. Cette façon d'aborder le problème est analogue à la recherche de bons codes en bloc avec la contrainte de se restreindre à des codes linéaires.

On dit qu'un empilement de sphères est un empilement de réseau, ou plus simplement un réseau si, quelles que soient deux sphères y appartenant, (l'une ayant pour centre le point \underline{P}_1 et l'autre le point \underline{P}_2), alors l'empilement contient aussi toutes les hypersphères centrées sur les points de la forme $a \underline{P}_1 + b \underline{P}_2$ où a et b sont des entiers quelconques. Plus formellement, on définit:

Définition (réseau).

Un empilement est appelé réseau s'il existe n vecteurs $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ linéairement indépendants dans l'espace euclidien à m dimensions, R^m , $m \geq n$, tels que les centres des sphères sont l'ensemble de tous les vecteurs $\underline{x} = \theta_1 \underline{v}_1 + \theta_2 \underline{v}_2 + \dots + \theta_n \underline{v}_n$ où les θ_i sont des entiers relatifs arbitraires.

Ainsi, un réseau réel Λ est tout simplement un ensemble discret (non dénombrable) de vecteurs dans l'espace euclidien R^m qui forme un groupe sous l'addition vectorielle ordinaire. Ce concept a pour origine historique la cristallographie.

Un réseau de rayon ρ est formé des centres $\{\underline{x}_1, \underline{x}_2, \dots\}$ vérifiant $d^2(\underline{x}_i, \underline{x}_j) \geq (2\rho)^2$, $i \neq j$. La distance euclidienne minimale entre deux points du réseau est donc $d_{\min} = 2\rho$.

Dans un empilement de type réseau, le nombre de contacts est commun à toutes les sphères: $\tau(\underline{x}) = \tau \forall \underline{x} \in \Lambda$.

Dans cette étude, on ne s'intéresse qu'à des empilements sur l'anneau des entiers relatifs. Leurs régions fondamentales sont des polytopes, c'est-à-dire des régions convexes de R^n limitées chacune par un nombre fini d'hyperplans. [CON&SLO 1982].

Définition (base d'un réseau - paralléloétope fondamental).

Les vecteurs $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ ci-dessus constituent une base du réseau. Le polytope (paralléloétope) formé de tous les points de l'espace euclidien

$$\theta_1 \underline{v}_1 + \theta_2 \underline{v}_2 + \dots + \theta_n \underline{v}_n \quad 0 \leq \theta_i < 1$$

est dit le "paralléloétope fondamental" du réseau.

Définition (dimension d'un réseau).

On appelle dimension d'un réseau Λ , notée $\dim \Lambda$, le nombre maximal de centres linéairement indépendants.

Définition (*matrice génératrice d'un réseau*).

Soit Λ un réseau de dimension n dans l'espace R^m , $n \leq m$. On appelle matrice génératrice du réseau une matrice M constituée des coordonnées d'une base $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$, $n \leq m$, soit

$$M = \begin{pmatrix} \underline{v}_1 \\ \underline{v}_2 \\ \vdots \\ \underline{v}_n \end{pmatrix}$$

Définition (*matrice de Gram*).

Soit une matrice génératrice M . On appelle $A := M M^t$ une matrice de Gram du réseau.

Définition (*déterminant d'un réseau*).

Le déterminant d'un réseau est le volume de son parallélotope fondamental, c'est-à-dire, $\det \Lambda = \det^{1/2} A$. Si M est carrée, alors $\det \Lambda = |\det M|$.

[Attention!: ce déterminant est quelquefois défini par $\det \Lambda = \det A$.]

La densité d'un réseau (fraction du volume de l'espace euclidien couverte par les sphères de l'empilement) est donc

$$\Delta = \frac{\text{volume d'une sphère inscrite}}{\text{volume d'une région fondamentale}} = \frac{V_n \rho^n}{\det \Lambda}.$$

Un paramètre d'un réseau très utilisé par les chimistes [BER 1969] est sa porosité, fraction du volume total qui n'est pas contenue dans les sphères de l'empilement: $\nabla := 1 - \Delta$.

Définition (*équivalence entre réseaux*).

On dit que deux réseaux Λ_1 et Λ_2 sont équivalents, ce que l'on note $\Lambda_1 \sim \Lambda_2$, si, et seulement si, ils ne diffèrent que par une rotation et un changement d'échelle.

Définition (*réseau dual*).

On dit qu'un réseau Λ^\perp est le dual d'un réseau Λ s'il est constitué de tous les points y du sous-espace de R^m engendré par $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ et tels que le produit scalaire $\underline{x} \cdot \underline{y} = \sum_{i=1}^n x_i y_i$ est un nombre entier quel que soit le point $\underline{x} \in \Lambda$ du réseau.

Les réseaux autoduaux sont tels que $\Lambda = \Lambda^\perp$, ce qui revient à $|\det \Lambda| = 1$. Mentionnons que les deux plus importants réseaux de la théorie des empilements, E_8 et Λ_{24} , sont des réseaux autoduaux.

Une région fondamentale d'un réseau quelconque Λ à n dimensions est définie comme une région de l'espace R^n qui ne contient qu'un seul point de chaque classe d'équivalence modulo Λ . Bien entendu, il y a plusieurs façons différentes de choisir une telle région, mais le volume qui en résulte est toujours le même. Il est appelé "volume fondamental" du réseau. Ainsi, la région fondamentale d'un réseau est en quelque sorte un motif qui pave tout l'espace euclidien de telle manière qu'un seul point du réseau se trouve dans chaque "pavé". De même qu'il y a différents choix possibles d'une base, il existe un grand nombre de régions fondamentales différentes, mais le volume $V(\Lambda)$ d'un "pavé" est déterminé de manière unique par Λ .

Autour de chaque point \underline{x} du réseau, on définit sa région (ou cellule) de Voronoi, $V(\underline{x})$, ensemble de tous les points de l'espace euclidien plus proches de ce point que de n'importe quel autre point du réseau. Plus exactement,

$$V(\underline{x}_i) = \{ \underline{x} \in R^n : ||\underline{x} - \underline{x}_i|| \leq ||\underline{x} - \underline{x}_j|| \quad \forall j \neq i \}$$

Puisque tous les points d'un réseau sont équivalents, il suffit de connaître la région de Voronoi entourant l'origine, $V(0)$. Supposant que les points du réseau sont des mots de code, la règle de décodage à vraisemblance maximale consiste à prendre les régions de Voronoi comme régions de décision.

Le volume d'une région de Voronoi peut être écrit en fonction des paramètres canoniques du réseau:

$$\text{vol } V(0) := \frac{V_n}{\Delta} = \frac{\rho^n}{\delta} = \det \Lambda$$

Un des paramètres importants d'une région R est son deuxième moment adimensionnel, noté $G_n(R)$ et défini par [CON&SLO 1982, FOR&WEI 1989]

$$G_n(R) := \frac{\int ||\underline{r}||^2 dV(\underline{r})}{n \left(\int dV \right)^{(1+2/n)}}$$

Ce moment, lié à la forme de la région, mesure l'efficacité avec laquelle elle remplit l'espace pour un volume donné (compacité).

Pour un cube à n dimensions, il vaut $G_n = 1/12$ quel que soit n . Il en est bien ainsi puisque le cube est le produit direct de segments linéaires. D'autre part, le deuxième moment d'une hypersphère à n dimensions (qui occupe le mieux l'espace avec une aire donnée) vaut

$$G_n = \frac{\Gamma(\frac{n}{2} + 1)^{2/n}}{(n+2)\pi} \rightarrow \frac{1}{2\pi e} \quad \text{quand } n \rightarrow \infty$$

Calderbank et Ozarov ont récemment étudié l'amélioration des performances qui résulte de l'utilisation de schémas où la distribution des points de la constellation n'est pas équiprobable [CAL&OZA 1990]. Il y est décrit comment s'approcher du gain de forme (*shape gain*) maximal quel que soit le nombre de dimensions considéré.

Définition (changement d'échelle d'un réseau).

Le changement d'échelle d'un réseau (ou d'un empilement) Λ par une constante réelle positive c a pour résultat le réseau $c\Lambda := \{c\underline{x} : \underline{x} \in \Lambda\}$ dont les principaux paramètres sont [SLO 1981]:

$$\rho(c\Lambda) = c\rho(\Lambda), \dim(c\Lambda) = \dim\Lambda, M(c\Lambda) = cM(\Lambda), \det(c\Lambda) = c^{\dim\Lambda} \det\Lambda,$$

$$\Delta(c\Lambda) = \Delta(\Lambda), \delta(c\Lambda) = \delta(\Lambda), \tau(c\underline{x}) = \tau(\underline{x}), \Theta_{c\Lambda}(z) = \Theta_{\Lambda}(c^2 z), \text{ et } (c\Lambda)^{\perp} = c^{-1} \Lambda^{\perp}.$$

Décomposition.

Soit L un réseau tel que $L = \bigcup_{j=0}^{t-1} (c_j + \Lambda)$. Les vecteurs c_j , $j = 0, 1, \dots, t-1$ sont appelés les représentants de classes latérales (*coset representatives*) de Λ dans L , et $t := |L|/|\Lambda| = \det L / \det \Lambda$ est dit l'index de Λ dans L .

Considérons par exemple $E_8 = (0^8 + D_8) \cup (1/2^8 + D_8)$. D_8 a pour représentants dans E_8 les vecteurs $(0,0,0,0,0,0,0,0)$ et $(1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 1/2)$ et son index y est deux.

La plupart des réseaux étudiés sont contenus dans leur dual [CON&SLO 1988,p.48], de sorte qu'il existe des représentants $\underline{c}_0, \underline{c}_1, \dots, \underline{c}_{d-1}$ tels que

$$\Lambda^{\perp} = \bigcup_{j=0}^{d-1} (\underline{c}_j + \Lambda)$$

d étant l'index de Λ dans Λ^{\perp} .

Un arrangement quelconque d'hypersphères qui ne se chevauchent pas (un empilement) ne suffit pas à couvrir tout l'espace euclidien, quel que soit le nombre n des dimensions avec $n > 1$.

Un autre problème classique fortement lié à l'empilement de sphères consiste à couvrir tout l'espace avec des hypersphères non disjointes. Dans ce cas, le but est de trouver, dans l'espace euclidien à n dimensions, le pavage le plus économique. Autrement dit, il s'agit maintenant de minimiser les régions de chevauchement en faisant un choix approprié des centres des hypersphères. Ceci est connu sous le nom de problème du "recouvrement" ou du "remplissage total" de l'espace.

Définition (rayon de recouvrement).

Le rayon de recouvrement d'un empilement (réseau) est défini par le *supremum* de la distance euclidienne minimale entre deux points quelconques [CON&SLO 1988]

$$r := \sup_{y \in R} \inf_{x \in \Lambda} d(x, y)$$

Bien évidemment, l'ensemble des sphères de rayon r centrées en chaque point d'un réseau suffit à recouvrir l'espace euclidien tout entier, propriété qui cesserait d'être vraie pour un rayon plus petit.

Les paramètres d'un recouvrement sont définis de la même façon que ceux d'un réseau. D'abord, l'épaisseur d'un recouvrement, Θ , définie par

$$\Theta := \frac{\text{volume d'une sphère circonscrite}}{\text{volume d'une région fondamentale}} = \frac{V_n r^n}{\det \Lambda}.$$

Il va de soi que $\Theta \geq 1$, alors que l'on a $\Delta \leq 1$ pour la densité d'un réseau.

On définit aussi l'épaisseur normalisée d'un recouvrement de façon analogue à celle de la densité de centres δ d'un réseau, soit $\theta := \Theta/V_n$.

Pour un empilement du type réseau ayant des cellules de Voronoi congruentes aux régions fondamentales, $V(\Lambda)$, le rayon de l'empilement ρ est celui de la sphère inscrite (*inradius*) dans $V(\Lambda)$ tandis que le rayon de recouvrement r est celui de la sphère circonscrite (*circumradius*) à $V(\Lambda)$.

Définition (*série θ*).

La série θ d'un réseau Λ est définie par

$$\Theta_{\Lambda}(z) := \sum_{x \in \Lambda} q^{|x|^2} = \sum_{m=0}^{\infty} N_m q^m$$

où $q = \exp(\pi i z)$, $i := \sqrt{-1}$.

Il suffira ici d'interpréter Θ_{Λ} comme une série formelle de puissances d'une indéterminée q , mais des investigations plus profondes devraient la considérer comme une fonction analytique de la variable complexe z , pourvu que $\text{IM}(z) > 0$. En outre, on remarque que la série Θ est déterminée par le réseau Λ , mais que l'inverse n'est pas vrai [CAL&SLO 1988,p.47].

On obtient la série θ du réseau dual Λ^{\perp} à partir de la série du réseau Λ selon:

$$\Theta_{\Lambda^{\perp}}(z) = (\det \Lambda)^{1/2} (i/z)^{n/2} \Theta_{\Lambda}(-1/z)$$

Cette identité peut être considérée comme une généralisation aux réseaux de l'identité de MacWilliams des codes en bloc linéaires [MacW&SLO 1977].

La détermination de la plus grande densité que peuvent atteindre des réseaux dans des espaces à grand nombre de dimensions reste un problème ouvert. Herman Minkowski démontrait en 1905 l'existence, pour toute dimension n , d'empilements de réseau dont la densité est supérieure à $\exp_2(-n)$ [CAS 1959]. Tout comme d'autres déjà mentionnées, sa démonstration n'est pas constructive et aucune famille de réseaux ayant cette densité n'a jamais été construite.

La borne inférieure de Minkowski sur la densité d'un réseau est

$$\Delta \geq \zeta(n) / \exp_2(n-1)$$

où $\zeta(n) := \sum_{k=1}^{\infty} k^{-n}$ est la fonction zéta de Riemann [ABRA&STE 1972].

Quand $n \rightarrow \infty$, alors $\log_2 \Delta \geq -n+1$.

Du côté des bornes supérieures, Blichfeldt démontrait en 1914 que, pour un nombre de dimensions arbitrairement grand, la densité d'un empilement ne pouvait excéder environ $\exp_2(-0,5n)$ [SLO 1984]. Le même résultat peut être obtenu par la borne de Rogers sur la densité de centres d'un réseau [LEBCH 1967],

$$\log \delta < \frac{n}{2} \log\left(\frac{n}{4\pi e}\right) + \frac{3}{2} \log n - \log \frac{e}{\sqrt{\pi}} + \frac{5,25 \log 2}{n+2,5}$$

Ainsi, les réseaux les plus denses dans l'espace euclidien R^n à nombre de dimensions assez élevé possèdent une densité satisfaisant à la double inégalité:

$$-n \leq \log_2 \Delta < -0,5n.$$

Plus récemment, G. Kabatiansky, V. Levenstein et V. Sidel'nikov ont montré que, pour tout nombre de dimensions assez grand, la densité d'un empilement de réseau ne peut dépasser environ $\exp_2(-0,599n)$ [CON&SLO 1988].

Barnes et Sloane ont présenté des constructions explicites d'empilements de réseaux dans l'espace R^n , $n \leq 98\,328$, dont la densité est d'environ $\Delta \approx \exp_2(-1,25n)$ [BAR&SLO 1983].

Les bornes sur l'épaisseur du meilleur recouvrement à nombre de dimensions n assez grand sont [CON&SLO 1988]:

$$\frac{n}{e\sqrt{e}} \leq \Theta < n \ln n + n \ln \ln n + 5n$$

Transformation orthogonale d'un réseau.

On note T une transformation orthogonale avec ou sans changement d'échelle.

L'application de T à un réseau Λ quelconque, $T\Lambda$, est encore un réseau. On ne s'intéresse ici qu'à un opérateur de rotation de 45° , noté R . (Bien entendu, cela définit un réseau équivalent.)

L'opérateur bidimensionnel de rotation est

$$R := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

tandis que pour n dimensions, n pair, il est donné par

$$R = \begin{pmatrix} R & \underline{0} & \underline{0} & \cdots & \underline{0} \\ \underline{0} & R & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{0} & R & \cdots & \underline{0} \\ \cdots & \cdots & \cdots & \cdots & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \cdots & R \end{pmatrix}$$

matrice $(n/2) \times (n/2)$ dont chaque entrée est une matrice 2×2 . Le réseau RZ^2 se déduit donc de Z^2 par une rotation de 45° et un changement d'échelle de $\sqrt{2}$. On voit aisément que $R^2 = 2I$, I étant la matrice identité à n dimensions.

Réseaux Z^n :

Les réseaux cubiques à n dimensions correspondent à la grille formée de tous les points dont les coordonnées sont des nombres entiers quelconques. Il s'agit d'un réseau auto-dual, $Z^n = (Z^n)'$.

$$Z^n := \{(x_1, x_2, \dots, x_n) : x_i \in Z\}.$$

$$\text{Sa série thêta est } \Theta_Z(z) = \sum_{m=-\infty}^{\infty} q^{m^2} = 1 + \sum_{m=1}^{\infty} 2q^{m^2}.$$

Réseaux A_n :

Pour $n \geq 1$, A_n est un réseau à n dimensions formé de tous les points (x_0, x_1, \dots, x_n) ayant des coordonnées entières dont la somme est nulle. Il est donc défini dans un espace à $m = n + 1$ dimensions. Le réseau le plus connu de cette famille est le réseau *hexagonal* à deux dimensions A_2 . On remarque encore que $A_2 \perp A_2^\perp$.

$$A_n := \{(x_0, x_1, x_2, \dots, x_n) \in Z^{n+1} : \sum x_i = 0\}.$$

Réseaux D_n :

Le réseau D_n à n dimensions est formé de tous les points (x_1, x_2, \dots, x_n) ayant des coordonnées entières dont la somme est un nombre pair. Ces réseaux sont aussi appelés "réseaux échiquier". En effet, un réseau de cette famille se déduit du réseau cubique en prenant les centres des sphères en quinconce sur le réseau, comme on dispose des cases

noires et blanches sur un échiquier. Si l'on colorie les nœuds d'un réseau cubique alternativement en rouge et en blanc, le réseau D_n correspondant est formé alors des points rouges. Le plus important réseau de cette famille est D_4 , connu sous le nom de réseau de Schlœfli.

$$D_n := \{(x_1, x_2, \dots, x_n) \in \mathbb{Z}^n : \sum x_i \text{ est pair}\}.$$

Réseau E_8 :

Découvert en 1900 par T. Gosset, ce réseau est formé de tous les points de l'espace euclidien (x_1, x_2, \dots, x_8) ayant des coordonnées de la forme $x_i \in \mathbb{Z}$ ou $x_i \in \mathbb{Z} + 1/2$ et dont la somme est un nombre pair.

Les vecteurs les plus proches de l'origine (vecteurs minimaux) sont 112 vecteurs du type $(\pm 1^2, 0^6)$ et 128 vecteurs du type $(\pm 1/2^8)$ ayant un nombre pair de signes négatifs. Le nombre total de sphères en contact est donc $\tau = 240$.

Réseau de Leech Λ_{24} :

En 1964, Leech montrait que le code (parfait) de Golay pouvait être incrusté dans un objet encore plus fascinant: le réseau de Leech [LEECH 1967]. Il est formé de tous les points de l'espace euclidien à 24 dimensions dont les coordonnées obéissent aux deux conditions:

1) Chaque point du réseau est, modulo 2, congruent à un mot du code de Golay étendu $(24, 12, 8)$.

2) La somme modulo 4 des coordonnées de n'importe quel point du réseau est nulle.

Les centres pairs les plus proches de l'origine ont deux coordonnées ± 4 , ou bien huit coordonnées ± 2 , toutes les autres étant égales à zéro. On les note $(\pm 4^2, 0^{22})$ et $(\pm 2^8, 0^{16})$: il y a 98 256 points avec des coordonnées de cette forme. Les centres impairs les plus proches de l'origine ont vingt-trois coordonnées ± 1 et une seule coordonnée ± 3 , soit $(\pm 1^{23}, \pm 3)$. Il y a 98 304 points avec des coordonnées de cette forme. Le nombre total de sphères en contact est donc $\tau = 196\,560$. Des détails sur l'histoire de la découverte de ce réseau peuvent être trouvés dans [THO 1983].

Enfin, mentionnons que des groupes peuvent être associés à la structure d'un réseau. En particulier, on appelle groupe d'automorphisme $\text{Aut}(\Lambda)$ d'un réseau Λ l'ensemble des transformations (rotations et réflexions) de l'espace qui laissent l'origine invariante et appliquent le réseau sur lui-même [CON&SLO 1988].

TABLEAU VII. PARAMETRES DES PRINCIPAUX RESEAUX
A FAIBLE NOMBRE DE DIMENSIONS.

| réseau | dim | det | τ | ρ | r |
|----------------|------------|--------------|-----------|--------------|---------------------------------|
| Z^n | n | 1 | $2n$ | $1/2$ | $\rho\sqrt{n}$ |
| A_n | $n=m-1$ | $\sqrt{n+1}$ | $n(n+1)$ | $1/\sqrt{2}$ | $\rho[\frac{2a(m-a)}{m}]^{1/2}$ |
| D_n | $n \geq 4$ | 2 | $2n(n-1)$ | $1/\sqrt{2}$ | $\rho\sqrt{n/2}$ |
| E_8 | 8 | 1 | 240 | $1/\sqrt{2}$ | $\rho\sqrt{2}$ |
| K_{12} | 12 | 27 | 756 | 1 | $\rho\sqrt{8/3}$ |
| Λ_{16} | 16 | 16 | 4320 | 1 | $\rho\sqrt{3}$ |
| Λ_{24} | 24 | 1 | 196560 | 1 | $\rho\sqrt{2}$ |

$$a := \lceil m/2 \rceil$$

TABLEAU VIII. PARAMETRES DES PRINCIPAUX RESEAUX
A FAIBLE NOMBRE DE DIMENSIONS (suite).

| réseau | Δ | δ | Θ | θ |
|----------------|-------------------------|---------------------|-----------------------------------|-------------------------------|
| Z^n | $V_n 2^{-n}$ | 2^{-n} | $V_n n^{n/2} 2^{-n}$ | $n^{n/2} 2^{-n}$ |
| A_n | $V_n m^{-1/2} 2^{-n/2}$ | $2^{-n/2} m^{-1/2}$ | $V_n m^{-(m+2)/2} [a(m-a)]^{n/2}$ | $m^{-(m+2)/2} [a(m-a)]^{n/2}$ |
| D_n | $V_n 2^{-(n+2)/2}$ | $2^{-(n+2)/2}$ | $V_n n^{n/2} 2^{-n+1}$ | $n^{n/2} 2^{-n+1}$ |
| E_n | $\pi^4/384$ | $1/16$ | $\pi^4/24$ | 1 |
| K_{12} | $\pi^6/19440$ | $1/27$ | $(8\pi/3)^6/19440$ | $8^6/3^9$ |
| Λ_{16} | $\pi^8/16 \cdot 8!$ | $1/16$ | $(3\pi)^8/16 \cdot 8!$ | $3^8/4^2$ |
| Λ_{24} | $\pi^{12}/12!$ | 1 | $(2\pi)^{12}/12!$ | 2^{12} |

Leurs séries thêta respectives sont données par [SLO 1981]:

$$\Theta_{Z^n}(z) = [\Theta_Z(z)]^n, \text{ où } \Theta_Z(z) = 1 + 2q + 2q^4 + 2q^{16} + \dots$$

$$\Theta_{D_n}(z) = 1 + 24q^2 + 96q^6 + 24q^8 + 144q^{10} + \dots$$

$$\Theta_{E_n}(z) = 1 + 240q^2 + 2160q^4 + 6720q^6 + \dots$$

$$\Theta_{K_{12}}(z) = 1 + 756q^4 + 4032p^6 + 20412q^8 + 60480q^{10} + \dots$$

$$\Theta_{\Lambda_{16}}(z) = 1 + 4320q^4 + 61440q^6 + 522720q^8 + \dots$$

$$\Theta_{\Lambda_{24}}(z) = 1 + 196560q^4 + 16773120q^6 + \dots$$

Dans le but de comparer différents réseaux n'ayant pas le même nombre de dimensions, il serait souhaitable de normaliser la mesure de densité à un même nombre de dimensions. Normalement, on le fait en choisissant la dimension 2 pour permettre une comparaison avec le réseau non codé Z^2 . Pour une vue plus globale, on peut considérer un réseau à n dimensions, ayant la densité D , que l'on rapporte à un nombre de dimensions n' , soit $D^{\pi'/n}$ (de Oliveira).

TABLEAU IX. DENSITE DES RESEAUX Z^2 , E_8 et Λ_{24} ,
RAPPORTEE A 2, 8 ET 24 DIMENSIONS.

| Λ | Δ | Δ à 2 dim. | Δ à 8 dim. | Δ à 24 dim. |
|----------------|------------------------|----------------------------|-----------------------------|---------------------------|
| Z^2 | $\frac{\pi}{4}$ | $\frac{\pi}{4}$ | $\frac{\pi^4}{4^4}$ | $\frac{\pi^{12}}{4^{12}}$ |
| E_8 | $\frac{\pi^4}{384}$ | $\frac{\pi}{384^{1/4}}$ | $\frac{\pi^4}{384}$ | $\frac{\pi^{12}}{384^3}$ |
| Λ_{24} | $\frac{\pi^{12}}{12!}$ | $\frac{\pi}{(12!)^{1/12}}$ | $\frac{\pi^4}{(12!)^{1/3}}$ | $\frac{\pi^{12}}{12!}$ |

Un tableau similaire pour la densité de centres est:

TABLEAU X. DENSITE DE CENTRES DES RESEAUX Z^2 , E_8 et Λ_{24} ,
RAPPORTEE A 2, 8 ET 24 DIMENSIONS.

| Λ | δ | δ à 2 dim. | δ à 8 dim. | δ à 24 dim. |
|----------------|----------|-------------------|-------------------|--------------------|
| Z^2 | 1/4 | 1/4 | 1/256 | 1/2 ²⁴ |
| E_8 | 1/16 | 1/2 | 1/16 | 1/2 ¹² |
| Λ_{24} | 1 | 1 | 1 | 1 |

De toute évidence, bien que l'agrandissement de l'espace conduise à une diminution de la densité, on s'aperçoit qu'il entraîne un accroissement de la *densité de centres* -

paramètre en rapport étroit avec le débit d'information.

Gersho et Lawrence ont bien saisi l'importance des réseaux comme technique de codage pour apporter des gains importants sur le canal gaussien [GER&LAW 1984]. Ils ont présenté une mise en œuvre très intéressante du codage/décodage à 4 et 8 dimensions et ont analysé les gains ainsi obtenus. Quoique leur notation ne soit pas la même, le premier système correspond exactement au réseau D_4 dont le gain théorique est 1,51 dB par rapport au système non codé ayant le même débit (le gain mesuré a été 1,2 dB). Le deuxième système correspond au réseau de Gosset E_8 dont le gain est 3,01 dB (le gain mesuré a été 2,4 dB). On trouve dans leur article le schéma d'un décodeur pratique pour des codes de réseau. Ils ont aussi indiqué la possibilité très attrayante d'utiliser le réseau de Leech. Enfin, ils ont vérifié expérimentalement que le codage par un réseau à 4 dimensions ne change en rien la densité spectrale en ligne, ce qui est un cas particulier d'un des résultats originaux de cette thèse: un code de réseau quelconque ne modifie pas le spectre du signal (voir l'annexe de la section III.2.).

Wilson et collaborateurs ont étudié la construction de jeux de signaux à 4 dimensions à l'aide d'une technique (*frequency-reuse*) qui utilise deux champs électromagnétiques avec polarisation orthogonale sur la même porteuse de façon à doubler l'efficacité spectrale [WIL *et al.* 1984]. Ils décrivent des modulations à 4 dimensions qui correspondent exactement aux réseaux Z_4 , A_4 et D_4 . Ils trouvent que les meilleures performances sont obtenues avec le réseau D_4 , comme on pouvait s'y attendre.

Le décodage de réseaux particuliers peut être mis en œuvre en exploitant des propriétés mathématiques intrinsèques du réseau pour en réduire la complexité. Bien des travaux ont été consacrés à ce sujet, ainsi pour les réseaux de Gosset [CON&SLO 1982, FOR 1988, SEC-deB 1989] et de Leech [CON&SLO 1986, FOR 1988, LAN&LON 1989, BE'E *et al.* 1989, FOR 1989]. Par exemple, la complexité du décodage du réseau de Leech a pu être réduite successivement de 56 000 opérations à 15 000, 10 000, puis 8 000 et finalement à moins de 2 000 opérations pour un décodeur sous-optimal ayant par rapport à l'optimalité une perte estimée à 0,1 dB!

Leeuwijn-Boullé et Belfiore ont cherché des constellations basées sur un empilement de sphères et bien adaptées aux canaux avec évanouissement de Rayleigh [LEE *et al.* 1991]. Dans l'espace à deux dimensions, ils ont trouvé une nouvelle constellation, optimale pour ce canal, qui y est donc meilleure que le réseau hexagonal A_2 optimal en l'absence de fluctuations. Il lui est équivalent, se déduisant de A_2 par une rotation de 45° et un changement d'échelle dont le facteur dépend du rapport signal à bruit moyen.

I.4.2.2 CODES DE CLASSES LATÉRALES.

"Trellis Codes are to lattice codes as convolutional are to block" G.D. FORNEY.

Forney et al. ont remarqué en 1984 que les techniques de modulation codée pouvaient être vues comme une combinaison, d'une part d'un codeur binaire classique (en bloc ou convolutif) dont les bits codés sortants sont utilisés pour sélectionner un des sous-ensembles d'une partition de la constellation, et d'autre part, de bits non codés additionnels utilisés pour choisir un seul point parmi tous ceux du sous-ensemble sélectionné [FOR et al. 1984]. Calderbank et Sloane ont envisagé la modulation codée du point de vue des réseaux et classes latérales, ce qui a été probablement la plus grande contribution conceptuelle apportée au codage-modulation depuis les travaux d'Ungerboeck. Dans leur remarquable article [CAL&SLO 1987], ils proposent une interprétation alternative (voire plus rigoureuse) de la technique de partition des ensembles d'Ungerboeck en termes de partitions de réseaux. En 1988, Forney a réussi à généraliser la théorie du codage-modulation en rassemblant tous les schémas connus sous un même formalisme mathématique: les codes de classes latérales (*coset codes*) [FOR 1988].

Définition (*classe latérale* ou *coset*).

Soit U une classe (abstraite) d'objets munie d'une opération binaire (abstraite) d'addition \oplus telle que $\langle U, \oplus \rangle$ est une classe fermée par rapport à l'addition. Soit C une sous-classe de U ($U \supseteq C$). Etant donné un élément u quelconque, on appelle l'ensemble $u + C := \{u \oplus x : x \in C\}$ une classe latérale (*coset*) de C .

On peut constater aisément que deux classes latérales sont disjointes à moins qu'elles ne coïncident [MacW&SLO 1977,p.15]. Géométriquement, les classes latérales correspondent à un pavage de l'ensemble C .

Calderbank et Sloane ont remarqué avec justesse qu'il ne s'agit pas à proprement parler d'une nouvelle façon de coder, mais plutôt d'une nouvelle manière (plus performante, cela va de soi) d'aborder la technique de partition des ensembles d'Ungerboeck. La partition de la constellation de signaux en sous-ensembles ayant des distances euclidiennes qui augmentent à mesure que l'on progresse dans la partition (c'est-à-dire que le niveau, ou la profondeur, de la décomposition en sous-ensembles augmente) correspond à une partition d'un réseau en les classes latérales d'un sous-réseau donné.

Les codes de classes latérales sont donc des séquences de *cosets* d'un sous-réseau Λ' dans la partition Λ/Λ' d'un réseau binaire Λ . Ainsi, une partie des bits non codés entrants est utilisée pour un codage convolutif généralisé dont la sortie désigne des classes latérales (*cosets*) du réseau Λ' tandis que l'autre partie est employée pour sélectionner un point dans les classes latérale ainsi choisies.

Un code de classes latérales est donc défini à partir des éléments suivants: d'abord, un réseau binaire Λ (se rapporter à la section précédente). Ensuite un sous-réseau Λ' ($\Lambda \supseteq \Lambda'$) lequel définit une partition Λ/Λ' (le groupe quotient) ayant $r=|\Lambda/\Lambda'|=2^{k+r}$ cosets de Λ' dans Λ . Enfin, un code correcteur d'erreur binaire de taux $k/(k+r)$. En résumé,

- Λ réseau (constellation)
- Λ' sous-réseau (sous-constellation)
- Λ/Λ' partition du jeu de signaux en $|\Lambda/\Lambda'|$ classes latérales (ordre de la partition).
- Un code binaire C.

Un code de classes latérales $C(\Lambda/\Lambda'; C)$ est l'ensemble de toutes les séquences de points dans le réseau (binaire) Λ , décomposé en classes latérales de Λ' , correspondant à la séquence de classes latérales que spécifie un train de bits codés selon un code binaire (classique) C. Cette généralisation peut inclure tous les schémas de codage-modulation connus, y compris les codes de réseau et les modulations codées en treillis MCT et en bloc MCB.

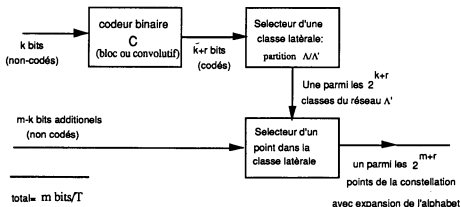


Figure 10. Structure générale d'un code de classes latérales $C(\Lambda/\Lambda';C)$.

Il y a 2^{k+r} cosets $\times 2^{m-k}$ points/coset, soit un total de 2^{m+r} points. Le débit de transmission d'information par deux dimensions est donc m bits/2D (bits/T).

Le sous-réseau Λ' provoque une partition de Λ (notée Λ/Λ') en classes d'équivalence modulo Λ' . Il y a $t = |\Lambda/\Lambda'|$ telles classes (l'index de Λ' dans Λ), valeur appelée l'ordre de la partition. Si l'on ne prend qu'un élément de chaque classe d'équivalence, alors on obtient un système de représentants des classes latérales de la partition, que l'on note $[\Lambda/\Lambda']$. Il est très important de souligner la relation suivante entre les volumes fondamentaux des deux réseaux [FOR 1988, Lemme 1]:

$$V(\Lambda') = |\Lambda/\Lambda'| V(\Lambda).$$

Décomposition en classes latérales (coset decomposition).

$$\Lambda = \bigcup_{j=0}^{t-1} \{ \xi_j + \Lambda' \}, \text{ ou en reprenant la notation de Forney, } \Lambda = \Lambda' + [\Lambda/\Lambda'].$$

Quelques exemples très naïfs illustrent la décomposition de réseaux simples en classes latérales:

| Partition | ordre de la partition | cosets | formule de décomposition |
|------------------------------|------------------------------------|---|---|
| $\mathbb{Z}/2\mathbb{Z}$ | $ \mathbb{Z}/2\mathbb{Z} = 2$ | $[\mathbb{Z}/2\mathbb{Z}] = \{(0), (1)\}$ | $\mathbb{Z}/2\mathbb{Z} = 2\mathbb{Z} \cup 2\mathbb{Z}+1$ |
| $\mathbb{Z}^2/R\mathbb{Z}^2$ | $ \mathbb{Z}^2/R\mathbb{Z}^2 = 2$ | $[\mathbb{Z}^2/R\mathbb{Z}^2] = \{(0,0), (1,0)\}$ | $\mathbb{Z}^2 = \{(0,0)+R\mathbb{Z}^2\} \cup \{(1,0)+R\mathbb{Z}^2\}$ |
| $\mathbb{Z}/m\mathbb{Z}$ | $ \mathbb{Z}/m\mathbb{Z} = m$ | $[\mathbb{Z}/m\mathbb{Z}] = \{0, 1, \dots, m\}$ | $\mathbb{Z}/m\mathbb{Z} = \bigcup_{j=0}^{m-1} (j + m\mathbb{Z})$ |

Le premier exemple fait une partition de l'ensemble des entiers en entiers pairs et impairs. La deuxième partition scinde la grille bidimensionnelle des entiers en deux sous-ensembles (en fait, deux sous-réseaux) qui correspondent au premier niveau de la partition d'Ungerboeck pour la modulation MAQ. Le troisième montre une partition des entiers dans m classes d'équivalences modulo m . Tout entier i peut être écrit d'une façon unique $i = km + j$, où km est un élément du réseau $m\mathbb{Z}$ des entiers multiples de m . Ainsi, j est le représentant de la classe d'équivalence de tous les entiers tels que $i = j \bmod m$.

Chaîne de partition

On appelle chaîne de partition en classes latérales et l'on note $\Lambda/\Lambda'/\Lambda''/\Lambda'''/\dots$ une suite de partitions provoquées par des réseaux emboîtés $\Lambda \supset \Lambda' \supset \Lambda'' \supset \Lambda''' \supset \dots$

Par exemple, la suite de réseaux $Z^4 \supset D_4 \supset RZ^4$ induit une chaîne de partition $Z^4/D_4/RZ^4$. La première partition, Z^4/D_4 , décompose le réseau Z^4 . Le réseau D_4 , quant à lui, il peut être décomposé selon la partition correspondant au groupe quotient D_4/RZ^4 .

Un exemple assez éclairant présenté par Forney consiste en la chaîne de partition $Z/2Z/4Z/\dots$ de sous-ensembles d'entiers. La chaîne de décomposition en classes latérales que l'on en tire, $Z = [Z/2Z] + [2Z/4Z] + [4Z/8Z] + \dots$, amène directement à la représentation binaire d'un nombre entier m arbitraire, $m = a_0 + 2a_1 + 4a_2 + \dots$

Décomposition en chaîne de classes latérales (Coset decomposition chain).

Un réseau peut être décomposé en fonction des classes latérales d'une chaîne de partition. Pour une chaîne $\Lambda/\Lambda'/\Lambda''/\Lambda'''/\dots$ donnée,

$$\Lambda = [\Lambda/\Lambda'] + [\Lambda'/\Lambda''] + [\Lambda''/\Lambda'''] + \dots$$

Des exemples intéressants de décomposition en chaîne de classes latérales peuvent être obtenus à partir des chaînes de partition suivantes:

$$Z / 2Z / 4Z / \dots$$

$$Z / RZ / 2Z / 2RZ / 4Z / \dots$$

$$Z^2 / RZ^2 / 2Z^2 / 2Z^2 / 2RZ^2 / 4Z^2 / \dots$$

$$Z^4 / D_4 / RZ^4 / RD_4 / 2Z^4 / 2D_4 / 2RZ^4 / \dots$$

Un réseau binaire est une généralisation des codes binaires en bloc. Forney a effectué la décomposition de la plupart des réseaux importants de dimension inférieure à 24. Cependant, il a remarqué l'existence de quelques réseaux indécomposables.

On résume ici des résultats de décomposition de réseaux binaires:

1°) Un réseau réel à n dimensions est un réseau binaire modulo 2 si, et seulement si, il est constitué de l'ensemble de tous les n -uplets congruents modulo 2 à l'un des mots d'un code en bloc linéaire binaire (n,k) , C .

$$\Lambda = 2Z^n + C \quad (\text{Formule clé de la décomposition}).$$

Cette formule explicite le réseau comme étant l'union de 2^k cosets de $2Z^n$

Exemples: Le réseau de Schläfli $D_4 = 2Z^4 + (4,3,2)$, ce que l'on notera dorénavant $D_4 \Leftrightarrow (4,3,2)$. On a alors:

$$Z^2 \Leftrightarrow (2,2,1), \quad RZ^2 \Leftrightarrow (2,1,2), \quad 2Z^2 \Leftrightarrow (2,0,\infty);$$

$$Z^4 \Leftrightarrow (4,4,1), \quad D_4 \Leftrightarrow (4,3,2), \quad RZ^4 \Leftrightarrow (4,2,2), \quad RD_4 \Leftrightarrow (4,1,4), \quad 2Z^4 \Leftrightarrow (4,0,\infty).$$

2°) Un réseau réel à n dimensions est un réseau binaire modulo 4 si, et seulement si, il est décomposable selon

$$\Lambda = 4Z^n + 2C_1 + C_0 \quad (\text{Formule clé de la décomposition}).$$

Exemple: Le réseau de Barnes-Wall $\Lambda_{16} = 4Z^{16} + 2(16,15,2) + (16,5,8)$.

Soit le code en treillis d'Ungerboeck basé sur un code convolutif dont le treillis a 4 états destiné à la transmission de 5 bits/2D et dont le gain fondamental est de 3,01 dB. La modulation non codée ayant le même débit d'information utilise une constellation MAQ à $2^5=32$ états. La modulation codée est fondée sur une constellation agrandie du type MAQ à 64 états. La technique de partition d'Ungerboeck interprétée en termes de classes latérales correspond au choix suivant:

- $\Lambda = Z^2$ le réseau bidimensionnel des entiers*,

- $\Lambda' = 2Z^2$ le réseau bidimensionnel des entiers pairs.

La partition Λ/Λ' est donc $Z^2/2Z^2$ dont l'ordre est $|Z^2/2Z^2| = 4$ comme le montre la figure 11 où les quatre classes latérales sont représentées.

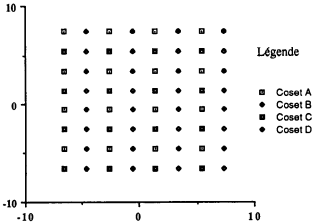


Figure 11. Une partition $Z^2/2Z^2$ sur un jeu de signaux à 64 états. L'ordre de la partition est $|Z^2/2Z^2| = 4$. Les quatre classes latérales sont représentées.

Lorsque l'on envisage les distances euclidiennes pertinentes à un tel processus de codage, deux types de distances interviennent:

$$d^2_1 := \min_{\substack{\mathbf{x} \in \Lambda' \\ \mathbf{x} \neq \mathbf{0}}} \|\mathbf{x}\|^2 = d^2_{\min}(\Lambda'),$$

et

$$d^2_2 := \min_{\mathbf{x} \neq \mathbf{y}} \|\mathbf{x} - \mathbf{y}\|^2 = d_H(C).$$

La première distance, $d^2_{\min}(\Lambda')$, est le carré de la distance euclidienne minimale entre deux points quelconques dans une même classe latérale de Λ' . Dans le diagramme de treillis du code, cela équivaut aux distances entre les transitions parallèles d'une même branche. La deuxième distance correspond quant à elle à la distance de Hamming minimale du code convolutif (ou en bloc), compte non tenu des branches parallèles ajoutées.

La distance euclidienne minimale d'un tel code de classes latérales est, bien évidemment, donnée par

$$d_{\min}^2(\mathbf{C}) = \text{MIN} (d_{\min}^2(\Lambda'), d_H(\mathbf{C})),$$

ou, en employant la notation de Calderbank et Sloane [CAL&SLO 1987] (à ceci près que l'on utilise le carré de la distance lui-même au lieu de le noter d)

$$d^2 = \text{MIN} (d_1^2, d_2^2).$$

Dans la partition de la figure précédente, le carré de la distance euclidienne minimale entre deux points d'une même classe latérale de $2Z^2$ est $d_2^2 := d_{\min}^2(\Lambda') = 4$. Le code convolutif est choisi quant à lui de façon que $d_1^2 := d_H(\mathbf{C}) = 5$. Le carré de la distance euclidienne minimale est donc $d_{\min}^2 = \text{MIN} (5, 4) = 4$. Le gain asymptotique (ou gain fondamental de Forney) est alors de 3,01 dB.

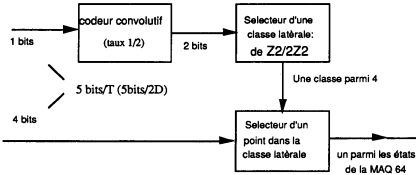


Figure 12. Code de classes latérales correspondant au code en treillis d'Ungerboeck basé sur un code convolutif avec treillis à 4 états, destiné à la transmission de 5 bits/2D et dont le gain fondamental est 3,01 dB.

Outre l'introduction d'une nouvelle technique pour la construction de codes MCT/MCB, Calderbank et Sloane ont aussi évalué les gains asymptotiques (fondamentaux) apportés par des codes conçus selon ces nouveaux principes. En reprenant la notation de Forney [FOR 1988], on écrit

$$m = |\mathcal{T}|^{2/n} = 2^{2(k+r)/n} \quad \text{et} \quad \rho_1 = k \log_2 a / \log_2 |\mathcal{T}| = \frac{k}{k+r} \log_2 a \text{ bits/coset};$$

$$\text{d'où } m^{\rho_1} = 2^{2k/n \log_2 a}.$$

Théorème (Calderbank-Sloane). Le gain asymptotique (en décibels) apporté par un code de réseau de dimension n est donné par

$$G(\Lambda) = 10 \log_{10} \frac{n+2}{3} \Delta^{2/n}$$

Théorème (Calderbank-Sloane). Le gain asymptotique (en décibels) apporté par un code de classes latérales est donné par

$$G(\mathcal{C}) = 10 \log_{10} \frac{n+2}{3} \frac{d^2}{d^2_2} \Delta^{2/n} 2^{2k/n \log_2 a}$$

Un point de vue personnel.

Il est possible d'interpréter la MCT, en faisant appel à la théorie des codes de classes latérales, comme étant une construction de réseau de dimension infinie.

$\mathcal{C}(\Lambda/\Lambda'; \mathcal{C})$ est un empilement de dimension infinie dont les paramètres sont:

- rayon de l'empilement:

$$\rho = d_{\min}(\mathcal{C})/2;$$

$$d^2_{\min}(\mathcal{C}) = \text{MIN}(d^2_{\min}(\Lambda'), d_H(\mathcal{C})).$$

- volume d'une région fondamentale (determinant de l'empilement):

$$V(\mathcal{C}) = \det \mathcal{C} = 2^{-k} V(\Lambda') = 2^{n-k} V(\Lambda);$$

$$V(\Lambda') = \det(\Lambda') \text{ et } V(\Lambda) = \det(\Lambda).$$

- densité de l'empilement:

$$\Delta(\mathcal{C}) = \frac{V_n \rho^n}{\det \mathcal{C}} = \frac{\rho^n}{V(\mathcal{C})}.$$

- densité de centres:

$$\delta(\mathcal{C}) = \frac{V_n \rho^n}{\det \mathcal{C}} = \frac{\rho^n}{V(\mathcal{C})} = \frac{1}{2^{-k}} \left[\frac{d_{\min}^2(\mathcal{C})}{d_{\min}^2(\Lambda')} \right]^{n/2} \delta(\Lambda').$$

- densité de centres rapporté à deux dimensions:

$$\delta^{2/n}(\mathcal{C}) = \frac{1}{2^{-2k/n}} \left[\frac{d_{\min}^2(\mathcal{C})}{d_{\min}^2(\Lambda')} \right] \delta^{2/n}(\Lambda').$$

Ainsi, les formules des gains fondamentaux (gains asymptotiques) selon Forney ou Calderbank-Sloane [FOR 1988, CAL&SLO 1987], soit:

$$\gamma_{\mathcal{C}}(\Lambda) := \frac{d_{\min}^2(\Lambda)}{V^{2/n}(\Lambda)} \quad \text{codes de réseau,}$$

$$\gamma_{\mathcal{C}}(\mathcal{C}) := \frac{d_{\min}^2(\mathcal{C})}{V^{2/n}(\mathcal{C})} \quad \text{codes de classes latérales,}$$

peuvent être réécrites ainsi d'une façon plus convenable (voir aussi sec.III.2):

$$\gamma_{\mathcal{C}}(\Lambda) := 4 \delta^{2/n}(\Lambda) = \delta^{2/n}(\Lambda) / \delta^{2/n}(\mathbb{Z}^n),$$

$$\gamma_{\mathcal{C}}(\mathcal{C}) := 4 \delta^{2/n}(\mathcal{C}) = \delta^{2/n}(\mathcal{C}) / \delta^{2/n}(\mathbb{Z}^n).$$

Qu'il s'agisse de codes de réseau ou de classes latérales, les relations précédentes mettent en évidence que le gain apporté par le codage est essentiellement dû à l'amélioration de la densité du jeu de signaux par rapport au système non codé \mathbb{Z}^n .

Ainsi, les gains asymptotiques donnés par Calderbank et Sloane (voir les théorèmes précédents) peuvent être séparés en contributions dues à la forme et au codage:

$$G(\Lambda) = 10 \log_{10} \gamma_s(\Lambda) \gamma_{\mathcal{C}}(\Lambda) = \gamma_s(\Lambda) + \gamma_{\mathcal{C}}(\Lambda) \quad (\text{dB}),$$

$$G(\mathcal{C}) = 10 \log_{10} \gamma_s(\mathcal{C}) \gamma_{\mathcal{C}}(\mathcal{C}) = \gamma_s(\mathcal{C}) + \gamma_{\mathcal{C}}(\mathcal{C}) \quad (\text{dB}).$$

Dans les deux cas, le gain de forme est donné par le gain d'une région sphérique par rapport au cube à n dimensions [FOR *et al* 1984, table III] [FOR&WEI 1989]:

$$\gamma_s(\Lambda) = \gamma_s(\mathcal{C}) = \frac{n+2}{12} V_n^{2/n} = \frac{n'+1}{6} \frac{\pi}{(n'!)^{1/n'}}, \quad n' := n/2.$$

Tous ces résultats coïncident avec les expressions des gains fondamentaux trouvées par Forney [FOR 1989].

TABLEAU XI.

GAIN THEORIQUE DU MEILLEUR RESEAU CONNU A PLUSIEURS DIMENSIONS.

| dimension | réseau | δ | $\gamma_c(\Lambda)$ | $\gamma_s(\Lambda)$ | gain (dB) |
|-----------|----------------|-----------------|---------------------|---------------------|-----------|
| 2 | A_2 | $1/(2\sqrt{3})$ | 0,62 | 0,20 | 0,82 |
| 4 | D_4 | 1/8 | 1,51 | 0,45 | 1,96 |
| 8 | E_8 | 1/16 | 3,01 | 0,73 | 3,74 |
| 12 | K_{12} | 1/27 | 3,63 | 0,88 | 4,51 |
| 16 | Λ_{16} | 1/16 | 4,51 | 0,98 | 5,49 |
| 24 | Λ_{24} | 1 | 6,02 | 1,10 | 7,12 |
| 48 | P_{48a} | $(3/2)^{24}$ | 7,78 | 1,26 | 9,04 |
| 64 | P_{64c} | 2^{22} | 8,09 | 1,31 | 9,40 |
| ∞ | -- | -- | ∞ | 1,33 | ∞ |

(comparer au Tableau XIV p.192 [CAL&SLO 1987] et au Tableau IV p.640 [FOR *et al.* 1984]).

TRAVAUX PERSONNELS

"My mind wanders around, and I conceive of different things day and night. Like a science-fiction writer, I'm thinking, *What if it were like this? or, Is there an interesting problem of this type?*, and I'm not caring whether someone is working on it or not".

C.E. SHANNON.

CHAPITRE II

CERTAINS ASPECTS DE LA MODULATION ET DU CODAGE

II.1. CONSTELLATIONS EN CROIX GÉNÉRALISÉES ET CANAL SECONDAIRE OPPORTUNISTE

RÉSUMÉ

On introduit un étiquetage binaire optimal des symboles d'une constellation carrée. Il permet une détection rapide pour les systèmes de modulation d'amplitude en quadrature (MAQ) au sens étroit. Ensuite, une famille de constellations en croix généralisées, à deux dimensions, est présentée ainsi que des bornes supérieure et inférieure de leur probabilité d'erreur par symbole sur un canal idéal à bande limitée, qui généralisent les bornes déjà connues pour la MAQ classique. L'application de ce schéma à un canal secondaire opportuniste est analysée et l'on montre comment des débits fractionnaires (en bits par signal à deux dimensions) peuvent être obtenus sur des systèmes MAQ généralisés. Ces schémas de communication sont comparés aux constellations multidimensionnelles généralisées récemment introduites par Forney et Wei.

ON GENERALIZED CROSS CONSTELLATIONS AND THE OPPORTUNISTIC SECONDARY CHANNEL

ABSTRACT

An optimum bits-to-symbol mapping for square constellations is introduced which allows a fast detection in narrow-sense QAM systems. Then, a family of two-dimensional generalized cross constellations is presented as well as upper and lower bounds on the symbol error probability over an ideal band-limited channel which generalizes those previously known for conventional QAM. The application of this scheme to the opportunistic secondary channel is analysed and it is shown how fractional rates (in bits per 2-dimensional signal) can be supported on 2D generalized QAM systems. These signaling schemes are compared with multidimensional generalized constellations recently proposed by Forney and Wei.

II.1.1. INTRODUCTION

Digital schemes with coherent amplitude and phase modulation have long been used in data communication systems [1]. A clear example of this use is the broad commercial success of Quadrature Amplitude Modulation (QAM) in both MODEMs on telephone lines and digital microwave radio links. Over the additive white Gaussian channel (AWGN), Shannon's capacity theorem asserts that the maximum achievable rate (in bits per dimension) with virtually error-free communication solely depends on the signal-to-noise ratio (SNR). However, it seems more interesting to consider a normalized rate in bits per two dimensional signal (*bits/2D*) rather than in bits per dimension. For a band-limited channel with a bandwidth of B hertz, the normalized rate exactly equals the spectral efficiency in (bit/s)/Hz of a system admitting a signaling rate of $2B$ dim/s [2].

The average noise power in the channel is $N = N_0 B$, N_0 being the noise one-sided power spectral density. Assuming that the average received signal power is constrained to be equal to S , then the capacity C of the Gaussian channel is given by

$$C = \log_2(1+S/N) \quad \text{bits/2D or (bits/s)/Hz.} \quad (1)$$

A more conservative estimate of the possible rate is the cut-off rate R_0 which is given for this channel by [3] :

$$R_0 = \log_2\left(1 + \frac{S}{2N}\right) \quad \text{bits/2D or (bits/s)/Hz.} \quad (2)$$

Let us consider the telephone channel as an example. Of course, other impairments beside the additive noise are present on the voiceband channel, but we assume that disturbances like nonlinear distortion, phase jitter, frequency offset, etc, are controlled on private special conditioned lines so that the additive Gaussian noise becomes the main impairment. For a signal-to-noise ratio of about 28 dB, the following limits are found: $R_0 \approx 8$ bits/2D and $C \approx 9$ bits/2D. The ceiling 19.2 kbits/s proposed in [4] corresponds exactly to the cutoff rate above when $B = 2,400$ Hz. At present, we have commercial MODEMs (with coding) working at 7 bits/2D or 7 (bits/s)/Hz, a rate fairly close to the cut-off rate. The rate increase from near the cut-off rate towards the capacity can be expected to be very hard. The usefulness of fractional rates in such a region becomes obvious. Still, transmitting a non-integral number of bits/2D is becoming common with the advent of coded multidimensional constellations, after Wei's paper [5]. In coded systems, a precise coding gain estimate involves its comparison with an uncoded system of same spectral efficiency, without sacrificing rate nor requiring more bandwidth [6]. Section 2 excepted, this paper is mainly concerned with transmitting at fractional rates and with uncoded modulation schemes able to support these rates. The aim of this work is not to find good (dense) constellations but rather to provide a better understanding about the error probability calculation and the evaluation of the rate supported by constellations.

The paper is organized as follows. First, square constellations are considered in section 2. An optimal bits-to-symbol assignment (2D Gray mapping) is presented which minimizes the bit error probability on AWGN channels. A fast detection algorithm is discussed which reduces the implementation complexity of uncoded QAM systems, particularly for large signal sets. Digital signaling methods using 2D generalized cross constellations (especially those with quadrilateral symmetry) are discussed in section 3. Both the average and peak power requirements of such constellations are evaluated. Bounds for the symbol error probability in the presence of AWGN are given. The opportunistic secondary channel is discussed in section 4. It is shown how to employ generalized QAM for transmitting a non-integral number of bits per two-dimension signal. These signaling schemes based on 2D generalized constellations are compared with multidimensional generalized constellations. The paper ends, in section 5, with concluding remarks.

II.1.2. A FAST DETECTION ALGORITHM FOR SQUARE QAM CONSTELLATIONS

In this section, we present a particular assignment, to be referred to as "2D Gray mapping", which is optimum in the sense that it results in a minimum bit error probability over the Gaussian channel.

It is well-known that no bits-to-symbol assignment exists such that a one-to-one correspondence between the Hamming (HD) and squared Euclidean distances (ED) results, except for 2- or 4-point constellations. Nevertheless, in the case of some particular constellations we can design a mapping where the minimum ED increases as the HD increases (the key word here is *minimum*). Figure 1 below exhibits such a mapping.

A $(4M)$ -point constellation is generated iteratively from an M -point one by substituting the points of the latter for a square constellation, to be referred to simply as a square. The initial constellation is the square whose points are labelled according to figure 1a. The process of iteratively labelling the points of the $(4M)$ -point constellation is described as follows.

The first $\log_2 M$ bits which label all the points of a given square are identical to the label of the point of the M -constellation it replaces. Now, the last two bits of the labels are assigned to the 4 points of each square according to one of the 8 equivalent Gray codings of a square, one of which is illustrated in figure 1a. The choice of the particular Gray coding used for a given square is such that each point bears the same last-2-bit label as its closest neighbour belonging to another square. Therefore, the Gray patterns chosen to determine the last two bits are symmetric with respect to the boundaries of the squares. Clearly, 4 different Gray patterns are used, depending on the labelling of the initial 4-point constellation: the initial pattern, the two ones which result from it by symmetry with respect to the x - and the y -axes, and the one resulting from the product of the two symmetries i.e., the original pattern rotated by π .

If the Gray number n_G is defined as the average number of erroneous bits per erroneous symbol [7], assuming only errors in favour of the closest points, it is apparent that $n_G = 1$ for these assignments (compare this with CCITT recommendation V32 [8]).

Let b and b' be two binary words (labels) assigned to the signal points \underline{P} and \underline{P}' , respectively. Let d_H (resp. d_E) denote the Hamming (resp. Euclidean) distance. Then, the two measures HD and ED are related according to the following inequalities:

$$\begin{array}{ll} d_H(b, b') = 1 & d_E(\underline{P}, \underline{P}') \geq \sqrt{1} d_0 \\ d_H(b, b') = 2 & d_E(\underline{P}, \underline{P}') \geq \sqrt{2} d_0 \\ \text{If } d_H(b, b') = 3 & \text{then } d_E(\underline{P}, \underline{P}') \geq \sqrt{5} d_0 \\ d_H(b, b') = 4 & d_E(\underline{P}, \underline{P}') \geq \sqrt{8} d_0 \\ \text{etc} & \text{etc} \end{array} \quad (3)$$

where d_0 is the minimum Euclidean distance between two constellation points.

It can readily be seen that the most likely error patterns result in less faulty detected bits than any other bits-to-symbol assignment, and consequently that the bit error probability is minimized (we assume maximum likelihood detection and equally likely symbols). Applications involving coding are beyond the scope of this paper and will be reported elsewhere. In the following, we propose a straightforward way to carry out the mapping/demapping in practical implementations. The bits-to-symbol conversion (and *vice-versa*) is performed with the help of several coordinate systems as indicated in figure 2. These coordinate systems clearly are symmetric with respect to each other as already discussed for the "squares" which enable to construct a $(4M)$ -point constellation from an M -point one, as discussed at the beginning of Section 2.

In many instances, it will be necessary to employ a coordinatewise vector product denoted by $*$ and defined to be

$$\underline{P}_1 * \underline{P}_2 = (x_1 x_2, y_1 y_2) \quad \text{where} \quad \underline{P}_i = (x_i, y_i) \in R^2 \quad i = 1, 2. \quad (4)$$

Bits-to-symbol mapping

Let us consider an $M = 2^m$ -point (m even) two-dimensional square constellation. A binary word \underline{b} of length m is assigned to each signal point as indicated in figure 1. Initially, each two consecutive bits are coupled as a 2-dimensional vector $\underline{b}_i = (b_{i1}, b_{i2})$, defining $m/2$ vectors in such a way that the binary word \underline{b} can be written as

$$\underline{b} = (\underline{b}_{m/2-1}, \underline{b}_{m/2-2}, \dots, \underline{b}_1, \underline{b}_0). \quad (5)$$

The conversion of the binary information into a signal point can be made according to the following procedure:

step 1: Generate a new word \underline{B} by replacing 0's by -1 in \underline{b} , thus defining

$$\underline{B} = (\underline{B}_{m/2-1}, \underline{B}_{m/2-2}, \dots, \underline{B}_1, \underline{B}_0). \quad (6)$$

step 2: The constellation point is given by

$$\underline{P} = \underline{B}_{m/2-1} 2^{m/2-1} + \dots + \underline{B}_{m/2-1} * \underline{B}_{m/2-2} * \dots * \underline{B}_1 2^1 + \underline{B}_{m/2-1} * \dots * \underline{B}_1 * \underline{B}_0 2^0. \quad (7)$$

where $*$ denotes the vector product introduced in (4).

In order to illustrate this method, we shall consider the $M = 64$ -point 2D constellation represented in figure 1c. As an example, let us find the constellation point \underline{P} assigned to a given binary word, say (100011):

we have $\underline{B} = (1 \ -1 \ -1 \ -1 \ 1 \ 1) = (\underline{B}_2, \underline{B}_1, \underline{B}_0)$ so that

$$\underline{P} = \underline{B}_2 2^2 + \underline{B}_2 * \underline{B}_1 2^1 + \underline{B}_2 * \underline{B}_1 * \underline{B}_0 2^0 = (4, -4) + (-2, 2) + (-1, 1) = (1, -1).$$

The interpretation of this assignment method is quite simple. The first two bits (10) select the IV-th quadrant of the 64-point constellation so it is reduced to a 16-point one with its origin displaced to the (4, -4) point. The next two bits (00) select the III-rd quadrant in the latter constellation, so it is reduced to a 4-point constellation with its origin displaced from (-2, 2) with respect to the (4, -4) point. This means that the origin was now moved to the point (-2, -2). Lastly, the bits (11) select the I-st quadrant, i.e., the (1, -1) signal point.

Symbol-to-bits demapping

Let $\text{sgn}(\cdot)$ be an operator defined over R^2 which indicates only the sign of the two coordinates, that is,

$$\text{sgn}(\underline{P}_i) \stackrel{\Delta}{=} (\text{sgn } x_i, \text{sgn } y_i), \quad \text{given } \underline{P}_i = (x_i, y_i) \in R^2. \quad (8)$$

In this section, \underline{B}_i will denote the 2-dimensional vector obtained by

$$\underline{B}_i = \text{sgn}(\underline{P}_i). \quad (9)$$

Before we start describing the demapping algorithm, we find it convenient to introduce $m/2$ special points defined by:

$$\underline{C}_i \stackrel{\Delta}{=} 2i(1, 1) \text{ for } i = 0, 1, 2, \dots, m/2-1. \quad (10)$$

These points are shown in the constellation drawn in figure 2.

Let \underline{P} be the received noisy vector. A fast detection can be performed by the following algorithm:

set $i = m/2-1$; $\underline{P}_i = \underline{P}$ (initial conditions)

step 1: $\underline{B}_i = \text{sgn}(\underline{P}_i)$;
if $i = 0$, then stop detection.

step 2: $\underline{P}_{i-1} = \underline{B}_i * \underline{P}_i - \underline{C}_i$;

$i = i - 1$;
go to step 1.

The detected binary word \underline{b} is the one which corresponds to \underline{B} .

An advantage of this technique in recovering digital information is that it does not require look-up tables. It is also interesting to mention that step 1 runs exactly $m/2$ times and that a partial detection of two bits occurs at each run. This implies that the information can be delivered to the sink (destination) before the end of the overall detection process, which can be advantageous for some very high-speed applications.

Once again we present an illustrative example using the constellation represented in figure 1c. Assuming that the received vector is $\underline{P} = (-3.2, 4.9)$, the algorithm above results in :

$$\begin{array}{ll} \underline{P}_2 = (-3.2, 4.9) & \underline{B}_2 = (-1, 1) \\ \underline{P}_1 = (-0.8, 0.9) & \text{therefore } \underline{B}_1 = (-1, 1) \\ \underline{P}_0 = (-1.2, -1.1) & \underline{B}_0 = (-1, -1) \end{array}$$

Thus $\underline{B} = (-11 \ -11 \ -1-1)$ and the detected binary sequence is (010100), which is the constellation point closest to \underline{P} . One can easily find the pair of axes in figure 2 which is the coordinate system relevant to the points \underline{P}_i . Finally, the following comment may be useful. In cases where one wishes to use a constellation with a minimum distance between points d_0 ($d_0 \neq 2$), it suffices to scale both the transmitted vector \underline{P} and the received one $\underline{P}_{m/2-1}$ by a factor $d_0/2$.

II.1.3. PERFORMANCE OF GENERALIZED CROSS CONSTELLATIONS

Let us consider the use of some multiple signal phase-amplitude schemes over an AWGN channel. It is established [9] that uncoded QAM performance with rectangular constellations is given by

$$P_e(2)_M = 1 - \left[1 - \left(1 - \frac{1}{\sqrt{M}} \right) \text{erfc} \left[\sqrt{\frac{3\gamma_{av}}{2(M-1)}} \right] \right]^2, \quad (11)$$

where

$P_e(2)_M$ is the error probability per 2D symbol,

$M = 2^m$ is the number of signal points, m even, and

$\gamma_{av} = E_{av}/N_0$ is the signal-to-noise ratio, E_{av} being the average energy per 2D symbol.

Accordingly, γ_{av} represents the ratio of the average energy to the noise spectral density, both per 2D. Dividing both the numerator and denominator by two results in the same ratio per dimension: $\gamma_{av} = 2E_N/N_0$, where $E_N = E_{av}/2$ is the average energy per dimension. Thus, γ_{av} is expressed in [Watts.s/dim]/[(Watts/Hz)/dim], so it is a dimensionless quantity. Moreover, assuming a signaling rate of $2B$ dim/s we find this

is equivalent to the ratio S/N usually found in the capacity formula. We claim this SNR measure is more suitable for channels in the bandwidth-limited region than E_b/N_0 which is more interesting for channels in the power-limited region.

Of course, we may neglect $1/\sqrt{M}$ if M is large enough, yielding

$$P_e(2)_M \leq 1 - [1 - \operatorname{erfc}(\sqrt{\frac{3\gamma_{av}}{2(M-1)}})]^2, \quad (12)$$

where the bound gets tighter and tighter as M increases.

Forney and Wei defined the figure of merit of a constellation C [10] as:

$$CFM(C) = \frac{\Delta}{P(C)} \frac{d_{\min}^2}{P(C)}, \quad (13)$$

where $P(C)$ is the average power of C . Instead, we think it is a bit more convenient to use

$$FM(C) = \frac{\Delta}{P(C)} \frac{d_{\min}^2/4}{P(C)} = \frac{\rho^2}{P(C)}, \quad (14)$$

which is the same except for a scale factor. For example, a QAM system with a square M -point constellation has $CFM(C) = 6/(M-1)$ (cf. [10], p.879) while $FM(C) = 3/[2(M-1)]$. The latter is exactly the multiplying factor for γ_{av} in the error probability expressions. As a matter of fact, all the expressions and/or bounds on the error probability over an AWGN channel (coded or uncoded systems) can be expressed as functions of $\frac{\rho^2}{P(C)} \gamma_{av}$. Additionally, this little modification does not change the asymptotic gain formula when coded constellations are used, since

$$g_{\text{asyp}} = \frac{CFM(C)_{\text{cod}}}{CFM(C)_{\text{unc}}} = \frac{FM(C)_{\text{cod}}}{FM(C)_{\text{unc}}} = \frac{(d_{\min}^2)_{\text{cod}}}{(d_{\min}^2)_{\text{unc}}} \frac{P(C)_{\text{unc}}}{P(C)_{\text{cod}}}, \quad (15)$$

as usual in coded-modulation.

We think this figure of merit is more natural. Moreover, it can also be extended to lattice coded systems [11]. So, we prefer to consider normalized constellations with respect to $d_{\min} = 2$ whereas Forney and Wei use $d_{\min} = 1$ as normalized value (our choice corresponds to $\rho = 1$ in lattice terminology). Hence, the normalized average power values printed in our tables have been calculated for constellations where $d_{\min} = 2$, i.e., the signal points were taken in an odd-integer $2Z^2 + (1,1)$ lattice.

Ordinarily, the performance of QAM systems is analysed at high signal-to-noise ratios ($\gamma_{av} \gg 2(M-1)/3$), in which case a reasonable approximation of the symbol error rate is

$$P_e(2)_M \approx 2\operatorname{erfc}(\sqrt{\frac{3\gamma_{av}}{2(M-1)}}) \approx \exp[-FM(C) \gamma_{av}]. \quad (16)$$

Let us now introduce a family of least-energy constellations, namely generalized cross

constellation, respecting a perfect quadrilateral symmetry and constrained to the fact that all their points belong to the odd-integer grid. Figure 3 shows some of these constellations. It can be noticed that all $(2k)^2$ -point constellations are square signal sets.

Generally, a signal set is expressed by:

$$s_j(t) = A_{cj} u(t) \cos(2\pi f_c t) - A_{sj} u(t) \sin(2\pi f_c t) \quad j = 1, 2, \dots, M, \quad (17)$$

where $u(t)$ is a narrow-band waveform,

(A_{cj}, A_{sj}) are coordinates in the 2D constellation,

and f_c is the carrier frequency.

Assuming a signaling rate of $1/T$ bauds, an information symbol is transmitted each T seconds and the QAM signal $s(t)$ is given by

$$s(t) = \sum_k a_{ck} u(t-kT) \cos(2\pi f_c t) - \sum_k a_{sk} u(t-kT) \sin(2\pi f_c t), \quad (18)$$

where (a_{ck}, a_{sk}) corresponds to the k -th transmitted signal $(A_{cj(k)}, A_{sj(k)})$.

The signal $s(t)$ can also be written in terms of its complex envelope:

$$s(t) = \text{Re} \sum_k P_k u(t-kT) \exp(j 2\pi f_c t), \quad (19)$$

where the complex $P_k = a_{ck} + j a_{sk}$ represents a constellation point.

Each signal $s_j(t)$ is associated with a 2D point $P_j = A_{cj} + jA_{sj} = (A_{cj}, A_{sj})$. The energy of a constellation point is $|P_j|^2 = A_{cj}^2 + A_{sj}^2$ and the average energy per 2D-point is $E(C) = E(|P_j|^2) = E(A_{cj}^2 + A_{sj}^2)$. The average energy $E(C)$ is often expressed as a normalized power $P(C)$ i.e., T has been normalized to unity so both terms are considered interchangeable.

It is supposed that the signals on phase and quadrature carriers can perfectly be separated by coherent detection, and that maximum likelihood detection is carried out by the receiver.

Let $P_e(i\text{-th D})$ denote the probability of an erroneous decision in the i -th dimension, $i = 1, 2$. For simplicity, we shall consider constellations with quadrilateral symmetry and furthermore such that

$$P_e(1\text{-st D}) = P_e(2\text{-nd D}) = P_e(1)_M, \quad (20)$$

where $P_e(1)_M$ is the error probability per dimension for the M -ary QAM system. Generally, detection must take account of all the signal dimensions because there is some dependence between signal coordinates (except for square constellations).

Let $P_e(1|P_i)$ denote the conditional probability of an erroneous decision per dimension given that the signal point P_i was transmitted. Obviously, $P_e(1)_M$ can be evaluated as:

$$P_e(1)_M = \sum_{i=0}^{M-1} P(P_i) P_e(1|P_i), \quad (21)$$

where $P(P_i)$ is the probability of P_i to be sent.

In an unbounded constellation, all the signal points have the same decision region so $P_e(1|P_i)$ does not depend on the taken point, that is, $P_e(1|P_i) = p_e$. Substituting this in (21), it follows that $P_e(1) = p_e$ for the boundless constellation ($M \rightarrow \infty$). For bounded constellations, this is an upper bound on the one dimensional error rate (since $P_e(1|P_i) \leq p_e$), that is

$$P_e(1)_M \leq P_e(1)_\infty = p_e. \quad (22)$$

Likewise, the 2D error rate (symbol error probability), $P_e(2)_M$, is also upper bounded by the symbol error probability for the unbounded constellation. In this case, all points have identical Voronoi regions so that

$$P_e(2)_M \leq P_e(2)_\infty = 1 - [1 - p_e]^2 \quad (23)$$

since the noises added to both components are statistically independent. Therefore, the error rate of such a signal set is essentially controlled by p_e .

We shall now investigate the error probability per dimension in an unbounded constellation over an ideal band-limited AWGN channel.

THEOREM 1.

The one-dimension error probability p_e of M -ary QAM systems based on 2D generalized uncoded constellations (border effect neglected) over an AWGN channel is given by $p_e = \text{erfc}(\sqrt{FM(C)}\gamma_{av})$ ■
proof.

The proof is straightforward and is based on a number of known results.

Defining the symbol autocorrelation function as $R_P(j) \stackrel{\Delta}{=} E(P_k P_{k-j}^*)$, it can be shown that the spectral power density of the signal $s(t)$ is given by (provided the quadrilateral symmetry):

$$S_s(f) = \frac{1}{4T} R_P(0) [|U(f - f_c)|^2 + |U(-f - f_c)|^2], \quad (24)$$

where $U(f)$ is the Fourier transform of $u(t)$ introduced in (18). Of course, $R_P(0) = E(|P_j|^2) = E(C)$.

The average power P_{av} of the QAM signal is found by integrating (24), that is,

$$P_{av} = \frac{E(C)}{4T} \int_{-\infty}^{+\infty} [|U(f - f_c)|^2 + |U(-f - f_c)|^2] df = \frac{E(C)}{2T} \xi_u, \quad (25)$$

where $\xi_u = \int_{-\infty}^{+\infty} |u(t)|^2 dt = \int_{-\infty}^{+\infty} |U(f)|^2 df$ is the energy of the narrow-band pulse $u(t)$.

Hence, the average energy of the QAM signal is

$$E_{av} = P_{av} T = E(C) \xi_u / 2. \quad (26)$$

As explained above, we are interested in the signal-to-noise ratio

$$\gamma_{av} = \frac{2E_{av}}{N_0} = P(C) \frac{\xi_u}{2N_0}. \quad (27)$$

We shall use the pass-band representation [12] of the random noise signal $n_w(t)$ in a bandwidth B Hz. The noise is assumed to be *white* with spectral power density

$$S_{n_w}(f) = \begin{cases} N_0/2 & |f| < B; \\ 0 & \text{otherwise.} \end{cases} \quad (28)$$

The narrow-band noise waveform can be expressed as:

$$n_w(t) = n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t), \quad (29)$$

where $n_c(t)$ and $n_s(t)$ are both low-pass noise signals band-limited to $B/2$ having as power spectral density

$$S_{n_c}(f) = S_{n_s}(f) = \begin{cases} N_0 & |f| < B/2; \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

Since the noise is assumed to be *additive*, the received signal is $s(t) + n_w(t)$ where the signal and the noise are given by (18) and (29), respectively. The received waveform is then processed: a coherent detection is performed followed by matched filtering and decision-making.

In order to achieve the maximum transmission rate we need use the fastest signaling rate. We know the bound $2B$ dim/s on the number of dimension per second accommodated by a band-limited channel [2]. In the case of 2D constellations, a symbol corresponds to 2 dimensions so the transmission speed must be less than B bauds = B symbol per second.

On the other hand, Nyquist's criterion asserts that a bandwidth $1/2T$ (base band) or $1/T$ (bandpass) is the minimum bandwidth needed for transmitting without intersymbol interference (ISI) at a rate $1/T$ symbol per second. These two approaches confirm that $T = 1/B$ represents the minimum interval where we can transmit without intersymbol interference on a band-limited channel of bandwidth B Hz.

The equivalent low-pass signals at the receiver can be written as

$$r_c(t) = \sum_k a_{ck} u(t-kT) + n_c(t) \quad (31.a)$$

$$r_s(t) = \sum_k a_{sk} u(t-kT) + n_s(t). \quad (31.b)$$

Hereafter, the index c or s will be dropped because the signal detection is identical in both the quadrature and phase components. We write simply:

$$r(t) = \sum_k a_k u(t-kT) + n(t). \quad (32)$$

The impulse response of the matched filter is $h(t) = u^*(-t)$, so that its output is

$$y(t) = \int_{-\infty}^{+\infty} r(\alpha) h(t-\alpha) d\alpha = \sum_k a_k \int_{-\infty}^{+\infty} u(\alpha-kT) u^*(\alpha-t) d\alpha + \int_{-\infty}^{+\infty} n(\alpha) u^*(\alpha-t) d\alpha. \quad (33)$$

The output of the matched filter at the sampling instant nT is the decision variable y_n :

$$y_n = \sum_k a_k \int_{-\infty}^{+\infty} u(\alpha - kT) u^*(\alpha - nT) d\alpha + \int_{-\infty}^{+\infty} n(\alpha) u^*(\alpha - nT) d\alpha. \quad (34)$$

It is known [12] that the response of a linear filter to a random Gaussian process is a Gaussian random variable, so

$$w_n = \int_{-\infty}^{+\infty} n(\alpha) u^*(\alpha - nT) d\alpha \approx N(0, N_0 \xi_u). \quad (35)$$

On the other hand, defining $p(t) \stackrel{\Delta}{=} u(t) \cdot u^*(-t)$, it follows that

$$y_n = a_n p(0) + \sum_{k, k \neq n} a_k p(nT - kT) + w_n. \quad (36)$$

The first term is the desired signal while the second and the third ones represent the intersymbol interference and additive Gaussian noise, respectively.

We assume $u(t)$ real and such that the waveform $p(t)$ obeys the first Nyquist criterion so the ISI vanishes. Therefore, we meet here the well known equally split filtering between the transmitter and the receiver.

Finally, we have

$$y_n = a_n \xi_u + w_n. \quad (37)$$

Alternatively, we can consider a "normalized" decision value $Y_n \stackrel{\Delta}{=} a_n + W_n$, where $W_n \approx N(0, \sigma^2 \stackrel{\Delta}{=} N_0/\xi_u)$. It is trivially verified that the one-dimensional error probability p_e is

$$p_e = \operatorname{erfc}\left(\sqrt{\frac{d_{\min}^2/4}{2\sigma^2}}\right). \quad (38)$$

Keeping Forney's notation, we rewrite (38) as

$$p_e = \operatorname{erfc}\left(\sqrt{(d_{\min}^2/4)\gamma_{av}/P(C)}\right) = \operatorname{erfc}\left(\sqrt{FM(C)\gamma_{av}}\right), \quad (39)$$

Q.E.D. ■

A more accurate estimate of the error probability must take into account the boundedness of the constellation. The exact performance calculation is complicated due to the shape of Voronoi regions concerning the outmost points. The reduction in the estimate of $P_e(2)_M$ depends on the number of outmost points and can be obtained as follows:

THEOREM 2

The performance of M -ary QAM systems with symmetrical generalized cross constellations satisfies the bounds $1 - [1 - 2\alpha_L p_e + \beta_L p_e^2] \leq P_e(2)_M \leq 1 - [1 - 2\alpha_U p_e + \beta_U p_e^2]$, where p_e is the one-dimensional error probability given by Theorem 1 and α, β are real numbers (less than unity) defined below. ■

proof.

Let Γ_I (resp. Γ_O) be the set of all inner (resp. outer) points and let $|\Gamma|$ denote the cardinality of the set Γ . Two kinds of outmost points, namely type I and type II, can be considered. First, the points that are outmost regarding just one dimension, and secondly those which are outmost regarding both dimensions. Consequently the set of outmost points, Γ_O , can be partitioned into two disjoint subsets, $\Gamma_O^I \cup \Gamma_O^{II}$.

Moreover, we notice the presence of a few ill-defined inner points occurring in the $[(2k)^2-4]$ -point constellations, but we neglect it in this analysis.

We shall suppose henceforth, as usual, that all signal points are equally likely *a priori*. Thus: $P(\Gamma) = |\Gamma|/M$. Certainly $|\Gamma_I| + |\Gamma_O| = M$, so that $P(\Gamma_I) + P(\Gamma_O) = 1$.

In order to circumvent the difficulty associated with the dependence between signal coordinates, we have bounded the error rate by taking rectangular decision regions different from the Voronoi regions. The following relations are then established:

$$P_e(2|P_i \in \Gamma_I) = 1 - [1 - p_\varepsilon]^2 \quad (40.a)$$

$$1 - [1 - p_\varepsilon] \cdot [1 - p_\varepsilon/2] \leq P_e(2|P_i \in \Gamma_O^I) \leq 1 - [1 - p_\varepsilon]^2 \quad (40.b)$$

$$1 - [1 - p_\varepsilon/2]^2 \leq P_e(2|P_i \in \Gamma_O^{II}) \leq 1 - [1 - p_\varepsilon] \cdot [1 - p_\varepsilon/2]. \quad (40.c)$$

The symbol error probability of QAM systems can clearly be written as:

$$\begin{aligned} P_e(2)_M &= \sum_{P_i \in \Gamma_I} P(P_i) P_e(2|P_i) + \\ &+ \sum_{P_i \in \Gamma_O^I} P(P_i) P_e(2|P_i) + \sum_{P_i \in \Gamma_O^{II}} P(P_i) P_e(2|P_i). \end{aligned} \quad (41)$$

Applying relations (40) to the equation above, the following bound is obtained:

$$P_e(2)_M \geq 1 - [P(\Gamma_I) [1 - p_\varepsilon]^2 + P(\Gamma_O^I) [1 - p_\varepsilon] [1 - p_\varepsilon/2] + P(\Gamma_O^{II}) [1 - p_\varepsilon/2]^2] \quad (42)$$

which after some computation yields the lower bound:

$$P_e(2)_M \geq 1 - [1 - 2\alpha_L p_\varepsilon + \beta_L p_\varepsilon^2], \quad (43)$$

where $\alpha_L = P(\Gamma_I) + 0.75 P(\Gamma_O^I) + 0.5 P(\Gamma_O^{II})$

and $\beta_L = P(\Gamma_I) + 0.5 P(\Gamma_O^I) + 0.25 P(\Gamma_O^{II})$.

Furthermore, we have

$$P_e(2)_M \leq 1 - [P(\Gamma_I) [1 - p_\varepsilon]^2 + P(\Gamma_O^I) [1 - p_\varepsilon]^2 + P(\Gamma_O^{II}) [1 - p_\varepsilon] [1 - p_\varepsilon/2]], \quad (44)$$

yielding the following upper bound:

$$P_e(2)_M \leq 1 - [1 - 2\alpha_U p_\varepsilon + \beta_U p_\varepsilon^2], \quad (45)$$

where $\alpha_U = P(\Gamma_I) + P(\Gamma_O^I) + 0.75 P(\Gamma_O^{II})$

and $\beta_U = P(\Gamma_I) + P(\Gamma_O^I) + 0.5 P(\Gamma_O^{II})$. Q.E.D. ■

It can be seen that M approaching infinity (i.e., an unbounded constellation) implies $\alpha_L, \alpha_U \rightarrow 1$ and $\beta_L, \beta_U \rightarrow 1$, so that the upper and lower bounds agree and $P_e(2)_M$ reduces to (23) as expected. Furthermore, an $M = 2^m$ -point (m even) rectangular

constellation can be found for which

$$P(\Gamma_O^H) = 4/M; \quad P(\Gamma_I) = 1 - \frac{4}{M}(\sqrt{M}-1) \quad \text{and} \quad P(\Gamma_O^L) = \frac{4}{M}(\sqrt{M}-2) \quad \text{so that} \quad \alpha_L = \beta_L^2 = 1 - \frac{1}{\sqrt{M}} \quad \text{and the lower bound is reached (see (11)).}$$

It can be conjectured that the symbol error probability of any 2D cross constellation is well fitted by a curve $P_e(2)_M = 1 - [1 - 2\alpha p_e + \beta p_e^2]$ for some α, β both less than, but close to, unity.

The performance evaluation of generalized cross constellations by Theorem 1 or 2 needs the normalized average power $P(C)$ of the constellations. These values are presented in table II where the average power is also expressed in dB. Values of the peak-to-average power ratio PAR are also available. For comparison purpose, the average power of minimum energy constellations constrained to the fact that all points belong to the odd-integer $2Z^2 + (1,1)$ lattice is also included. Indeed, cross constellations are not always optimal, but they are attractive at least from a theoretical viewpoint. Constellations of minimum energy have normalized power denoted by $P(C)_{MIN}$ which is easily computed by taking the M least-energy points in Table I.

TABLE I.
Energy Distribution in the $2Z^2 + (1,1)$ lattice.

| Shell | N. of points | Energy per 2D | Shell | N. of points | Energy per 2D |
|-------|--------------|---------------|-------|--------------|---------------|
| 1 | 4 | 2 | 17 | 16 | 170 |
| 2 | 8 | 10 | 18 | 8 | 178 |
| 3 | 4 | 18 | 19 | 8 | 194 |
| 4 | 8 | 26 | 20 | 8 | 202 |
| 5 | 8 | 34 | 21 | 8 | 218 |
| 6 | 12 | 50 | 22 | 8 | 226 |
| 7 | 8 | 58 | 23 | 8 | 234 |
| 8 | 8 | 74 | 24 | 4 | 242 |
| 9 | 8 | 82 | 25 | 16 | 250 |
| 10 | 8 | 90 | 26 | 8 | 274 |
| 11 | 4 | 98 | 27 | 16 | 290 |
| 12 | 8 | 106 | 28 | 8 | 298 |
| 13 | 8 | 122 | 29 | 8 | 306 |
| 14 | 16 | 130 | 30 | 8 | 314 |
| 15 | 8 | 146 | 31 | 12 | 338 |
| 16 | 4 | 162 | 32 | 8 | 346 |

TABLE II.
Parameters of some Generalized Cross Constellations.

| M | P(C) | dB | PAR | P(C) MIN | dB | PAR |
|-----|---------|---------|--------|----------|---------|--------|
| 4 | 2.0000 | 3.0103 | 1.0000 | 2.0000 | 3.0103 | 1.0000 |
| 8 | 10.0000 | 10.0000 | 1.0000 | 6.0000 | 7.7815 | 1.6666 |
| 12 | 7.3333 | 8.6530 | 1.3636 | 7.3333 | 8.6530 | 1.3636 |
| 16 | 10.0000 | 10.0000 | 1.8000 | 10.0000 | 10.0000 | 1.8000 |
| 20 | 14.8000 | 11.7026 | 1.7567 | 13.2000 | 11.2057 | 1.9696 |
| 24 | 15.3333 | 11.8564 | 1.6956 | 15.3333 | 11.8564 | 1.6956 |
| 28 | 20.2857 | 13.0719 | 1.6760 | 18.0000 | 12.5527 | 1.8888 |
| 32 | 20.0000 | 13.0103 | 1.7000 | 20.0000 | 13.0103 | 1.7000 |
| 36 | 23.3333 | 13.6798 | 2.1428 | 23.3333 | 13.6798 | 2.1428 |
| 40 | 26.0000 | 14.1497 | 1.9230 | 26.0000 | 14.1497 | 1.9230 |
| 44 | 28.1818 | 14.4997 | 1.7741 | 28.1818 | 14.4997 | 1.7741 |
| 48 | 31.3333 | 14.9601 | 1.8510 | 30.6666 | 14.8667 | 1.8913 |
| 52 | 32.7692 | 15.1547 | 1.7699 | 32.7692 | 15.1547 | 1.7699 |
| 56 | 37.4285 | 15.7320 | 1.9770 | 35.7143 | 15.5284 | 2.0720 |
| 60 | 38.2666 | 15.8282 | 1.9337 | 38.2666 | 15.8282 | 1.9338 |
| 64 | 42.0000 | 16.2325 | 2.3333 | 41.0000 | 16.1278 | 2.0000 |
| 68 | 43.4117 | 16.3761 | 1.8888 | 43.4118 | 16.3761 | 1.8888 |
| 72 | 46.4444 | 16.6693 | 2.1100 | 46.0000 | 16.6276 | 1.9565 |
| 76 | 48.3157 | 16.8409 | 1.8627 | 48.3158 | 16.8409 | 1.8627 |
| 80 | 50.8000 | 17.0586 | 1.9291 | 50.8000 | 17.0586 | 1.9291 |
| 84 | 53.8095 | 17.3086 | 1.9699 | 53.4286 | 17.2777 | 1.9839 |
| 88 | 55.8181 | 17.4678 | 1.8990 | 55.8181 | 17.4678 | 1.8990 |
| 92 | 60.4347 | 17.8129 | 2.1510 | 58.6957 | 17.6861 | 2.0785 |
| 96 | 62.0000 | 17.9239 | 2.0967 | 61.3333 | 17.8770 | 1.9891 |
| 100 | 66.0000 | 18.1954 | 2.4545 | 64.0800 | 18.0672 | 2.0287 |
| 104 | 66.6153 | 18.2357 | 1.9515 | 66.6154 | 18.2357 | 1.9515 |
| 108 | 70.1481 | 18.4602 | 2.3093 | 68.9630 | 18.3862 | 1.8850 |
| 112 | 71.1428 | 18.5213 | 1.8273 | 71.1429 | 18.5213 | 1.8273 |
| 116 | 74.2758 | 18.7085 | 2.1810 | 73.7241 | 18.6761 | 1.9803 |
| 120 | 76.1333 | 18.8157 | 1.9176 | 76.1333 | 18.8157 | 1.9176 |
| 124 | 78.9032 | 18.9709 | 2.0531 | 78.9032 | 18.9709 | 2.0531 |
| 128 | 82.0000 | 19.1381 | 2.0731 | 81.7500 | 19.1249 | 2.0795 |

TABLE II. (Continued)

| M | P(C) | dB | PAR | P(C) MIN | dB | PAR |
|-----|---------|---------|--------|----------|---------|--------|
| 132 | 84.4242 | 19.2647 | 2.0136 | 84.4242 | 19.2647 | 2.0136 |
| 136 | 89.0588 | 19.4968 | 2.2681 | 86.9412 | 19.3923 | 1.9555 |
| 140 | 91.1428 | 19.5972 | 2.2163 | 89.3143 | 19.5092 | 1.9033 |
| 144 | 95.3333 | 19.7924 | 2.5384 | 91.7777 | 19.6274 | 1.9394 |
| 148 | 95.4054 | 19.7957 | 2.1172 | 94.1081 | 19.7363 | 1.8914 |
| 152 | 99.2631 | 19.9679 | 2.4379 | 96.7368 | 19.8559 | 2.0054 |
| 156 | 99.6410 | 19.9844 | 2.0272 | 99.2308 | 19.9665 | 1.9550 |
| 160 | 103.200 | 20.1368 | 2.3449 | 101.800 | 20.0775 | 1.9842 |
| 164 | 104.243 | 20.1805 | 1.9377 | 104.244 | 20.1805 | 1.9377 |
| 168 | 107.523 | 20.3150 | 2.2506 | 106.952 | 20.2919 | 2.0382 |
| 172 | 109.534 | 20.3955 | 1.9902 | 109.535 | 20.3955 | 1.9902 |
| 176 | 112.545 | 20.5133 | 2.1502 | 112.182 | 20.4992 | 2.0145 |
| 180 | 115.777 | 20.6363 | 2.1593 | 114.711 | 20.5961 | 1.9701 |
| 184 | 118.521 | 20.7380 | 2.1093 | 117.304 | 20.6931 | 1.9948 |
| 188 | 123.191 | 20.9058 | 2.3540 | 119.787 | 20.7841 | 1.9534 |
| 192 | 125.666 | 20.9922 | 2.3076 | 122.333 | 20.8754 | 1.9782 |
| 196 | 130.000 | 21.1394 | 2.6000 | 124.939 | 20.9670 | 2.0009 |
| 200 | 129.680 | 21.1287 | 2.2362 | 127.440 | 21.0531 | 1.9617 |
| 204 | 133.764 | 21.2634 | 2.5268 | 129.843 | 21.1342 | 1.9254 |
| 208 | 133.692 | 21.2611 | 2.1691 | 132.154 | 21.2108 | 1.8917 |
| 212 | 137.547 | 21.3845 | 2.4573 | 134.830 | 21.2979 | 2.0321 |
| 216 | 138.000 | 21.3988 | 2.1014 | 137.407 | 21.3801 | 1.9940 |
| 220 | 141.636 | 21.5117 | 2.3863 | 140.182 | 21.4669 | 2.0687 |
| 224 | 142.857 | 21.5490 | 2.0300 | 142.857 | 21.5490 | 2.0300 |
| 228 | 146.280 | 21.6519 | 2.3106 | 145.439 | 21.6268 | 1.9939 |
| 232 | 148.482 | 21.7168 | 2.0608 | 147.931 | 21.7006 | 1.9603 |
| 236 | 151.694 | 21.8097 | 2.2281 | 150.475 | 21.7746 | 1.9804 |
| 240 | 155.066 | 21.9052 | 2.2312 | 152.933 | 21.8450 | 1.9485 |
| 244 | 158.065 | 21.9884 | 2.1889 | 155.443 | 21.9157 | 1.9685 |
| 248 | 162.774 | 22.1159 | 2.4205 | 157.871 | 21.9830 | 1.9382 |
| 252 | 165.555 | 22.1894 | 2.3798 | 160.349 | 22.0507 | 1.9582 |
| 256 | 170.000 | 22.3045 | 2.6470 | 162.750 | 22.1152 | 1.9293 |

In order to check the validity of our results, the performance of a 12-ary QAM system in the presence of AWGN (generated by the polar method [13]) was evaluated by Monte Carlo simulation. Figure 4 represents computer simulation results of the symbol error probability versus signal-to-noise ratio. The least square regression analysis shows that the symbol error rate fits the $1 - [1 - 2 \cdot 0.706p_e + 0.475p_e^2]$ curve with a correlation coefficient $r = 0.999$ and a standard error of estimation $s = 0.013$ for $n = 18$ data points. The upper and lower bounds were also drawn for comparison (solid lines). It can be seen that the simulated performance closely agrees with the results predicted from Theorem 2.

Finally, we apply these bounds as to evaluate the performance of conventional cross constellation, namely 32-QAM and 128-QAM. Results are presented in table III.

TABLE III.
Error Performance of 32-QAM and 128-QAM Uncoded Signaling Schemes.

| M | $P(\Gamma_I)$ | $P(\Gamma_O')$ | $P(\Gamma_O'')$ | α_L | β_L | α_U | β_U |
|-----|---------------|----------------|-----------------|------------|-----------|------------|-----------|
| 32 | 1/2 | 1/4 | 1/4 | 0.8125 | 0.6875 | 0.9375 | 0.8750 |
| 128 | 11/16 | 4/16 | 1/16 | 0.9065 | 0.8281 | 0.9843 | 0.9687 |

II.1.4. OPPORTUNISTIC SECONDARY CHANNEL

The generalized cross constellations described in the previous section can be regarded as a way to convey a non-integral number of bits per symbol (or bits per 2D). In a recent paper [10], Forney and Wei discussed the use of generalized multidimensional constellations and their ability of supporting an opportunistic secondary channel. In earlier applications, this technique was conceived as a means to create an opportunity to transmit additional side information (like internal signaling and control data) [14]. More generally, it can be used for increasing the information bit rate by a small fractional number, i.e., as a way to introduce an additional rate in generalized constellations. We intend here to make a precise estimate of this potential additional normalized rate.

Given an M -point cross constellation, a starting (or basic) constellation is defined to be the largest 2^m -point constellation with size smaller than M . For the sake of simplicity, we show as an example the 7-point constellation drawn in figure 5. The

starting constellation (shaded points) has only 4 points like fig.1a. Each new additional point is associated with a point in the basic constellation in such a way that both have the same label. An opportunistic bit (written into brackets in fig.7) is appended which takes value 0 or 1 according to whether it belongs or not to the starting constellation, respectively.

We say that the direct channel was used when the label of the transmitted signal point have no opportunistic bit, otherwise we say that the opportunistic channel was selected. In figure 5, assuming the source bits to be equally likely, we have

$$P(\text{select opportunistic channel}) = P(\{(00,01,11)\}) = 3/4,$$

$$P(\text{select direct channel}) = P(\{(10)\}) = 1/4.$$

In cases where the direct channel is employed, only 2 information bits are transmitted while 3 information bits are conveyed when the opportunistic channel is selected. Thus, the average transmitted number of information bits per two-dimension is $\beta = \frac{1}{4} \cdot 2 + \frac{3}{4} \cdot 3 = 2.75$ bits/2D. This process can be generalized as follows. Let us denote by M the number of points in a 2D cross constellation. Let m be an integer such that $2^m \leq M < 2^{m+1}$. It can be readily verified that $m = \lfloor \log_2 M \rfloor$ is such an integer, where the floor function $\lfloor x \rfloor$ denotes the largest integer no greater than x , as usual. Of course, an integer W exists, $0 \leq W < 2^m$, such that

$$M = 2^m + W. \quad (46)$$

Here, W represents the number of points appended to the basic constellation. In the following, we generalize the use of the opportunistic secondary channel by the same reasoning as in the previous example. An amount of m bits are transmitted per symbol (2D) when the direct channel is used. In contrast, $m + 1$ bits (per 2D) are transmitted when the opportunistic secondary channel is selected.

Let $P_{D.C}$ and $P_{O.C}$ be the probability of selecting the direct channel and the opportunistic channel, respectively. The average transmission rate (in information bits per 2D) is given by

$$\beta = P_{D.C} \cdot m + P_{O.C} \cdot (m+1) \quad \text{bits/2D}. \quad (47)$$

For equally likely source bits, it follows that

$$\beta = \frac{2^m - W}{2^m} \cdot m + \frac{W}{2^m} \cdot (m+1) = m + \frac{W}{2^m}. \quad (48)$$

Since $0 \leq W < 2^m$, we can use the binary decomposition to write $W = \sum_{i=0}^{m-1} d_i 2^i$. Using Forney and Wei's notation, the normalized rate β in bits per 2-dimension is

$$\beta = \sum_{i=0}^m \frac{d_i 2^i}{2^m}, \quad (49)$$

where we have defined $d_m = \Delta$.

Now, we shall investigate an explicit relationship between the normalized bit rate β and the constellation size M . To begin with, let us consider a new function $(\log)_2^\dagger$ to be referred to as *rational logarithm* and defined for any positive integer M by

$$(\log)_2^\dagger M \stackrel{\Delta}{=} \left\lfloor \log_2 M \right\rfloor + \left\lfloor \frac{M}{2^{\left\lfloor \log_2 M \right\rfloor}} - 1 \right\rfloor. \quad (50)$$

Combining equations (48) and (46) and remembering that $m = \left\lfloor \log_2 M \right\rfloor$, it follows that

$$\beta = (\log)_2^\dagger M. \quad (51)$$

Bounds on the $(\log)_2^\dagger$ function are given by the following theorem:

THEOREM 3.

For any integer $M > 0$, the $(\log)_2^\dagger$ function verifies:

$$\left\lfloor \log_2 M \right\rfloor \leq (\log)_2^\dagger M \leq \log_2 M, \quad (52)$$

with equality iff $M = 2^m$ for some integer m . ■

proof.

The proof of the lower bound in (52) is immediate: we have

$$(\log)_2^\dagger M - \left\lfloor \log_2 M \right\rfloor = \frac{M}{2^{\left\lfloor \log_2 M \right\rfloor}} - 1 = 2^{\log_2 M - \left\lfloor \log_2 M \right\rfloor} - 1 \geq 0. \quad (53)$$

The upper bound can be proved in the following way: we can write from the $(\log)_2^\dagger$ definition:

$$\log_2 M - (\log)_2^\dagger M = \log_2 M - \left\lfloor \log_2 M \right\rfloor + 1 - 2^{\log_2 M - \left\lfloor \log_2 M \right\rfloor}. \quad (54)$$

On the other hand, one has $\left\lfloor \log_2 M \right\rfloor \leq \log_2 M < \left\lfloor \log_2 M \right\rfloor + 1$, so that

$$0 \leq \log_2 M - \left\lfloor \log_2 M \right\rfloor < 1. \quad (55)$$

Let us deal with the real function: $f(x) \stackrel{\Delta}{=} x + 1 - 2^x \quad -\infty < x < +\infty$ which satisfies the property:

$$f(x) \geq 0 \quad \text{for } 0 \leq x < 1. \quad (56)$$

Therefore, taking $x = \log_2 M - \left\lfloor \log_2 M \right\rfloor$, the claimed inequality $\log_2 M - (\log)_2^\dagger M \geq 0$ follows from (56) and (54). At last, we mention it is trivial to verify the conditions "iff" about equalities. Q.E.D. ■

We have evaluated the normalized rate β in bits/2D supported by the 2D signaling

schemes just described. A comparison between the multidimensional constellation schemes proposed by Forney and Wei (cf. [10], table II, p.891) and those suggested here is presented in Table IV. According to our point of view, an advantage of the 2D signaling schemes is to allow a more "realistic" comparison with the coded modulation systems at fractional spectral efficiencies. We think this could be more suitable than the baseline criterion proposed by Forney and Wei. Indeed, the practical applications are rather complicated owing to the problems caused by the variable length of the transmitted binary word in this kind of signaling systems [14].

TABLE IV.
Comparison between 2D and Multidimensional Uncoded Signaling Schemes.

| 2D Generalized Constellations | | | | | | Multidimensional Constellations | | | | | |
|-------------------------------|---------|---|--------|-------|------|---------------------------------|---------|----|--------|-------|------|
| M | β | N | P(C) | dB | PAR | $ C_2 $ | β | N | P(C) | dB | PAR |
| 128 | 7.0000 | 2 | 81.75 | 19.12 | 2.07 | 128 | 7.0000 | 2 | 81.37 | 19.09 | 2.07 |
| 144 | 7.1250 | 2 | 91.78 | 19.63 | 1.94 | 144 | 7.1250 | 16 | 87.24 | 19.39 | 2.10 |
| 160 | 7.2500 | 2 | 101.80 | 20.08 | 1.98 | 160 | 7.2500 | 8 | 94.24 | 19.72 | 2.16 |
| 180 | 7.4062 | 2 | 114.71 | 20.60 | 1.97 | 180 | 7.3750 | 16 | 101.80 | 20.06 | 2.25 |
| 192 | 7.5000 | 2 | 122.33 | 20.88 | 1.98 | 192 | 7.5000 | 4 | 112.04 | 20.47 | 2.18 |
| 216 | 7.6875 | 2 | 137.41 | 21.38 | 1.99 | 216 | 7.6250 | 16 | 121.20 | 20.81 | 2.27 |
| 240 | 7.8750 | 2 | 152.93 | 21.85 | 1.95 | 240 | 7.7500 | 8 | 132.12 | 21.19 | 2.31 |
| 256 | 8.0000 | 2 | 162.75 | 22.12 | 1.93 | 256 | 8.0000 | 2 | 162.76 | 22.10 | 2.00 |
| 270 | 8.0547 | 2 | 171.90 | 22.35 | 2.01 | 270 | 7.8750 | 16 | 143.72 | 21.55 | 2.39 |

II.1.5. CONCLUDING REMARKS

We presented an attractive mapping for rectangular constellations as well as a fast detection algorithm which seems to be interesting for practical implementation of QAM at high-speed data transmission. In particular, it would be of interest to high-capacity digital microwave radio systems. We obtained also upper and lower bounds of the symbol error probability for 2D generalized cross constellations in the presence of band-limited AWGN, and in particular, for conventional cross constellations ($M = 2^m$, m odd). The ability of 2D cross constellations in transmitting fractional rates was investigated making use of the opportunistic secondary channel. A closed-form expression of the normalized rate (in bits per 2D) as a function of the constellation size was obtained. These signaling schemes are suggested as a reference for performance comparison in coded-modulation systems with fractional normalized rates.

ACKNOWLEDGEMENTS

The first author wishes to thank his colleague Mr. W. Zhang for a number of stimulating discussions. He also thanks the partial support from the Association of Research and Teaching on Communications (ARECOM-FRANCE).

REFERENCES

- [1] C.N. Campopiano and B.G. Glazer, A Coherent Digital Amplitude and Phase Modulation Scheme, *IRE Trans. Commun. Syst.*, vol. CS10, n° 1, Mar. 1962, pp. 90-95
- [2] A.D. Wyner, The Capacity of the Band-limited Channel, *Bell System Tech. J.*, vol. 45, Mar. 1965, pp. 359-371
- [3] J.M. Wozencraft and I.M. Jacobs, Principles of Communication Engineering, NY:Wiley, 1965
- [4] G.D. Forney Jr., R.G. Gallager, G.R. Longstaff and S.U. Qureshi, Efficient Modulation for Band-Limited Channels, *IEEE J. Select. Areas Commun.*, vol. SAC 2, n° 5, Sept. 1984, pp. 632-646
- [5] L.F. Wei, Trellis-Coded Modulation With Multidimensional Constellations, *IEEE Trans. Info. Theory*, vol. IT 33, n° 4, July 1987, pp. 483-501
- [6] G. Ungerboeck, Trellis-Coded Modulation with Redundant Signal Sets. part I: introduction, *IEEE Commun. Mag.*, vol. 25, n° 2, Feb. 1987, pp. 12-21
- [7] J.-C. Bie, D. Duponteil and J.-C. Imbeaux, Éléments de communications numériques, Paris: Dunod, 1986, Tome I, p. 183
- [8] ***, Famille de modems à deux fils fonctionnant en duplex à des débits binaires allant jusqu'à 9 600 bit/s pour usage sur le réseau téléphonique général avec commutation et sur les circuits loués de type téléphonique, CCITT, Livre bleu, Genève, 1989, Tome VIII, fascicule VIII.1, Recommandation V.32, pp. 234-241
- [9] J.G. Proakis, Digital Communications, NY:McGraw-Hill, 1986
- [10] G.D. Forney Jr. and L.F. Wei, Multidimensional Constellations - Part I: Introduction, Figures of Merit, and Generalized Cross Constellations, *IEEE J. Select. Areas Commun.*, vol. SAC 7, n° 6, Aug. 1989, pp. 877-892
- [11] H. Magalhães de Oliveira and G. Battail, A Capacity Theorem for Lattice Codes on Gaussian Channels, proc. *IEEE/SBT Int. Symposium Telecommun.*, Sept. 5-9, 1990, Rio de Janeiro, Brazil, pp. 1.2.1-1.2.5
- [12] W.B. Davenport Jr. and W.L. Root, Random Signals and Noise, NY:McGraw-Hill, 1958
- [13] D.E. Knuth, The art of computer programming, Reading, Mass.: Addison-Wesley, 2-nd ed., 1973
- [14] R.D. Gitlin, H.K. Thapar and J.J. Werner, An InBand Coding Method for the Transmission of Secondary Data, ICC'88, Philadelphia, pp. 3.1.1-3.1.5

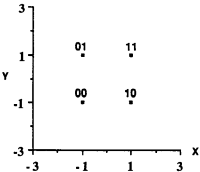


Fig. 1a.

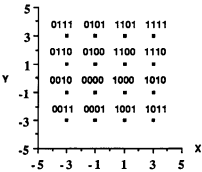


Fig. 1b.

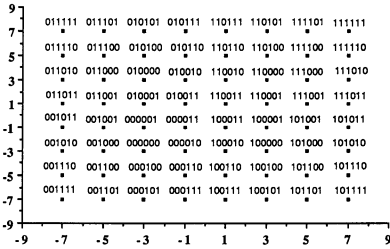


Fig. 1c.

Figure 1. This figure illustrates the iterative construction and bits-to-symbol assignments of a $(4M)$ -point constellation from an M -point one. Figure 1a is the starting 4-point constellation with Gray coding bits-to-symbol assignment. The 16-point constellation of figure 1b results from replacing each point of the previous one by a 4-point constellation; the bits-to-symbol assignment used is described in the text. The next step of this process i.e., the 64-point constellation, is illustrated in figure 1c.

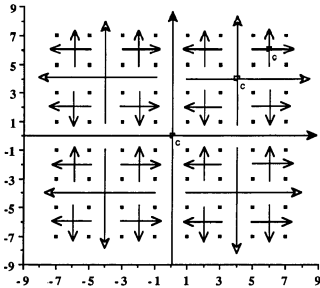


Figure 2. Coordinate systems which are used for both the bits-to-symbol conversion and the fast detection of received data when square constellations and bits-to-symbol assignement of figure 1 are used.

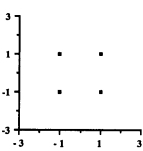


Figure 3a.

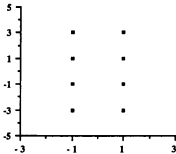


Figure 3b.

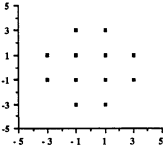


Figure 3c.

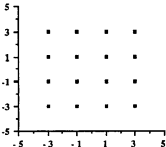


Figure 3d.

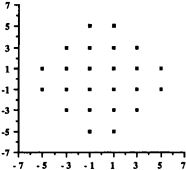


Figure 3e.

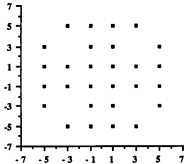


Figure 3f.

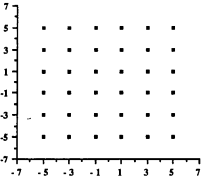


Figure 3g.

Figure 3. Generalized cross constellations with quadrilateral symmetry. Their sizes are 4 (3a), 8 (3b), 12 (3c), 16 (3d), 24 (3e), 28 (3f), and 36 (3g).

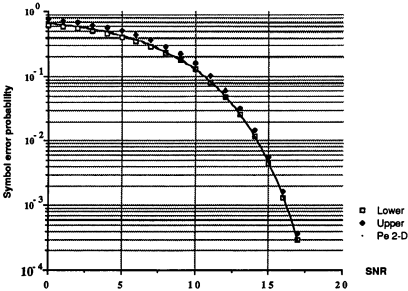


Figure 4. Bounds on the symbol error probability for a 12-point constellation and the corresponding simulated error rate.

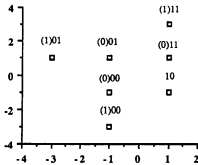


Figure 5. 7-point generalized constellation with opportunistic channel. The starting (or basic) constellation is the square 4-point one with shaded points. 3 more points are appended to 3 of its points. A third "opportunistic" bit (in brackets) is appended to their labels to tell whether the corresponding point belongs or not to the starting constellation.

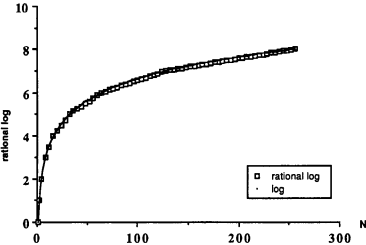


Figure 6. Rational log function and its relation with the log function.

II.2. PROBLEMES LIES A LA RECUPERATION DE PORTEUSE (MCT):

MODULATION DIFFERENTIELLEMENT COHERENTE.

Les problèmes de synchronisation constituent un chapitre intéressant de l'étude de la modulation codée. La récupération de la fréquence de rythme peut être réalisée de manière classique, mais la récupération de porteuse pose quelques problèmes. Ungerboeck [UNG 1982] observa que l'expansion de l'alphabet rend le jeu de signaux plus sensible aux erreurs de phase. Il remarqua, pour la MDP codée, l'existence d'un seuil de l'écart de phase au delà duquel les décisions du décodeur cessent d'être corrélées avec les signaux reçus; c'est alors le bruit qui pilote la boucle de récupération de la porteuse. Quoique le système puisse finir par se resynchroniser, ce comportement peut durer un certain temps. En outre, la détection cohérente est très difficile à mettre en œuvre dans bien des cas, tels que la communication radio mobile.

Lorenzelli *et al.* ont étudié les effets d'une synchronisation imparfaite de phase sur les systèmes MDP 4 non codée et MDP 8 codée en treillis [LOR *et al.* 1987]. Un système où la phase est récupérée à l'aide d'une horloge est analysé et comparé avec la récupération de porteuse par des circuits de multiplication de fréquence. Un algorithme séquentiel pour la détection différentielle de systèmes de MCT sur une modulation MDP est proposé par Makrakis et Feher [MAK&FEH 1988].

Edbauer considère l'application de la modulation codée aux canaux avec évanouissement. Il propose l'adaptation du décodage différentiellement cohérent à un code en treillis en construisant un code MCT sur une modulation différentielle 8 DMDP avec entrelacement [EDB 1987, EDB 1989a, EDB 1989b].

McLane et collaborateurs ont étudié une combinaison de codes en treillis avec une modulation différentielle par déplacement de phase (DMDP), dans le cadre du programme canadien de communications mobile par satellite [McL *et al.* 1988]. Le système travaille sur un canal à évanouissement de Rice et est essentiellement conçu pour la transmission de la voix, et des taux d'erreur de l'ordre de 10^{-3} suffisent à cette application.

Torck, Kawas-Kaleh et collaborateurs ont récemment abordé le décodage et l'estimation de phase conjoints pour la modulation codée, à l'aide d'une version modifiée de l'algorithme de Viterbi [TOR *et al.* 1991].

Le problème de la récupération de porteuse en présence de gigue de phase et d'écart de fréquence dans les systèmes de modulation MAQ à haute efficacité (> 6 bits/s/Hz) a été considéré par Cupo et Gitlin [CUP&GIT 1989]. Ils proposent un système adaptatif où la boucle d'asservissement qui commande l'oscillateur local utilise des filtres à réponse impulsive finie (RIF) ou infinie (RII), ce qui améliore la fiabilité de la réception par rapport aux boucles à verrouillage de phase classiques. Ces structures évitent l'antagonisme, inhérent à la plupart des boucles de récupération de porteuse, entre l'élargissement de la bande (favorable à la poursuite rapide de la phase) et son rétrécissement (pour minimiser le bruit à la sortie).

Nous avons abordé ici les performances de la modulation codée en treillis en présence d'un écart de phase. Nous nous sommes intéressé particulièrement aux problèmes de synchronisation de phase de la porteuse quand l'oscillateur local présente un déphasage par rapport à l'oscillateur de l'émetteur. Kawas Kaleh a suggéré de combiner le décodage différentiellement cohérent avec la modulation codée, comme solution pratique au cas de canaux où il est difficile de maintenir la cohérence de phase. L'introduction d'une phase aléatoire inconnue oblige à modifier la règle de décodage. On pourrait envisager, par exemple, d'utiliser encore l'algorithme de Viterbi [VIT 1967, VIT&OMU 1982], mais la distance euclidienne maximale cesse d'être le critère pertinent et le problème du choix d'un code convenable se pose.

Dans un premier temps, des constellations de signaux ayant tous la même énergie, en particulier les signaux MDP codés, ont seules été envisagées. Comme on verra par la suite, le décodage non cohérent optimal conduit à une métrique non additive. La complexité, même dans les cas les plus simples, rend le décodage exhaustif pratiquement impossible. Une approximation intéressante a consisté à ne retenir qu'un seul terme de l'estimation de la phase. La règle de décodage obtenue dans ce cas est fonction d'une combinaison entre un symbole s_k et son prédécesseur s_{k-1} . Le train $\{s_k\}$ subit donc un traitement (codage différentiel) dont le résultat est un train $\{s'_k\}$. En considérant la modulation différentielle comme une opération portant sur des symboles de canal complexes, on se ramène par un changement de variable à une règle de décision qui a la même expression que dans le cas cohérent. On a donc fait l'hypothèse que les suites de symboles *après* codage différentiel s'_k devaient avoir la distance euclidienne libre la plus grande possible, c'est-à-dire que les suites $\{s'_k\}$ doivent appartenir à un code d'Ungerboeck.

La conception du codeur devient très simple si on l'envisage sous la forme d'un codeur d'Ungerboeck suivi d'un codeur différentiel inverse. Ce système a été simulé en Pascal sur le VAX, dans le cas d'un code d'Ungerboeck à 4 états, et les courbes du taux d'erreur en fonction du rapport signal-à-bruit ont été tracées. Le gain (asymptotique) du système différentiel MDP 8 codé par rapport au système différentiel MDP 4 non codé est d'environ 3 dB pour des rapports signal à bruit élevés.

Mais les suites ainsi émises n'appartiennent pas à un code d'Ungerboeck et on peut chercher à concevoir un système en se donnant la contrainte que les suites émises appartiennent à un tel code. On aboutit alors à un code où le nombre d'états est le double de celui du code d'Ungerboeck, par scission de chacun des nœuds du treillis. Ainsi, les treillis à 4 et 8 états correspondant aux MDP 8 / MCT à 2 et 4 états, respectivement, ont été trouvés. La structure de corps sur les symboles du canal joue dans cette procédure un rôle important et les codes convolutifs obtenus ne sont pas linéaires. La simulation du code à 8 états a montré que les résultats étaient peu différents des précédents.

Encore qu'il se justifie par un raisonnement séduisant, l'entrelacement s'avère trompeur. Du strict point de vue de la théorie de l'information, le bruit le plus nuisible est complètement aléatoire. Si le bruit possède une structure quelconque, elle peut être exploitée au bénéfice de l'information utile. Un entrelaceur s'efforce donc soigneusement de convertir un bruit peu nuisible en un autre qui l'est davantage! Bien qu'il s'agisse d'une technique erronée au niveau philosophique, elle est très utile dans toutes les applications pratiques.

II.2.1. INTRODUCTION.

The performance analyses of coded modulation schemes are often carried out under the assumption of coherent demodulation. However, a practical problem which occurs in digital communications is the difficulty to extract a good carrier reference from the received signal. Thus, trellis-coded modulation schemes - which are known to be very efficient - could be made more robust by providing a differentially coherent detection in order to suppress such a requirement of carrier recovery (Fig. 1).

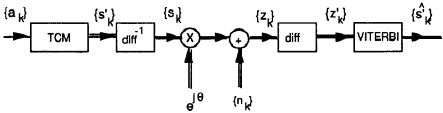


Figure 1. Block diagram of TCM system with differential detection.

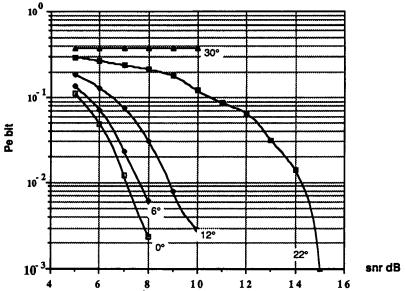


Figure 2. Sensitivity of bit error probability to carrier phase drift for 4-state 8 PSK/TCM Ungerboeck code.

Let us assume a binary input $\{a_k\}$ applied to a 2^V -state binary convolutional encoding (constraint length $v+1$) followed by a bits to symbol mapping (modulation).

Of course, $s_k = f(a_k, a_{k-1}, \dots, a_{k-v})$ and $s_{k-1} = f(a_{k-1}, a_{k-2}, \dots, a_{k-v-1})$, so that

$$s_k s_{k-1}^* = F(a_k, a_{k-1}, a_{k-2}, \dots, a_{k-v-1}).$$

An approach to Trellis-Coded Modulation with differential detection could therefore be the following: the decoder could use a 2^{V+1} -state trellis instead of the 2^V -state one needed in the coherent decoding owing to the correlation between consecutive channel symbols. This is equivalent, in some way, to enhance the memory of the encoder. We shall consider the project of such a trellis in the next section.

II.2.2. DESIGN OF TRELLIS-CODED M-ary DPSK MODULATION.

Let the channel symbols (phasor notation, $j = \sqrt{-1}$) be

$$s_k(i) = \sqrt{2E_s} e^{j\phi_k(i)}; \quad \phi_k(i) \in \Phi := \left\{ \frac{2\pi}{2M} i; i=0,1,\dots,2M-1 \right\},$$

where E_s is the energy per symbol and $\phi_k(i)$ is the phase.

Without loss of generality, we shall consider normalized signals, that is $\sqrt{2E_s}=1$.

The differential operation is represented by

$$s'_k(i) = s_k(i) s_{k-1}(i)^* = e^{j\phi_k(i)} e^{-j\phi_{k-1}(i)} = e^{j(\phi_k(i) - \phi_{k-1}(i))} \quad \text{where } \phi_k(i) - \phi_{k-1}(i) \in \Phi.$$

The differential encoding have some interesting properties. First, it does not change the shape of signals, that is, $s'_k(i)$ is a signal of the same kind as $s_k(i)$. Second, it is a reversible transformation in the sense that there exists an inverse operation such that, given a differentially coded sequence $\{s'_k(i)\}$, it is possible to recover the corresponding sequence $\{s_k(i)\}$, provided there is no noise.

It is straightforward to derive the sequence $\{s_k(i)\}$ from the sequence $\{s'_k(i)\}$. Clearly,

$$s_k = s'_k s_{k-1} \quad \text{because } s^*_{k-1} s_{k-1} = 1.$$

In view of the fact we consider an AWGN channel, it seems reasonable to use channel sequences belonging to an Ungerboeck code which have the largest squared free Euclidean distance. However, under the assumption that differential operation is required at the receiver, the trellis code we need will not be an Ungerboeck code. More precisely, we are looking for a trellis code such that, after differential encoding of its sequences, it results in coded sequences belonging to an Ungerboeck code (e.g. Fig.4a). Although we can use directly the related Ungerboeck code at the transmitter, we should know the "inner" code because the trellis to be used in the Viterbi decoder corresponds to that code. As said before, such a trellis code has twice the number of states of the "associate" Ungerboeck code.

Notice that differential operation on channel symbols involves multiplication and complex conjugate. For each $0 \leq i < 2M$ (expanded signal set) define channel symbols by $s(i) := \exp(2\pi\sqrt{-1} \frac{i}{2M})$. The set $S := \{s(i) \mid 0 \leq i < 2M\}$ is an Abelian group with respect to the multiplication $s(i).s(j) = s(i+j)$ $0 \leq i, j < 2M$, where $i+j$ is taken mod $2M$.

Clearly, the unity element is $s(0)$, $s(0).s(i) = s(i)$, and the inverse of an element $s(i)$ is $s(i)^*$, $s(i).s(i)^* = s(0)$, where $s(i)^* = s(2M-i)$ $0 \leq i < 2M$, $2M-i$ being taken mod $2M$.

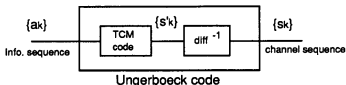


Figure 3. Block representation of the "differential" trellis code U' associated with an Ungerboeck code U .

• 4-state 8-DPSK/TCM

For simplicity, we first introduce as an example the 4-state 8-DPSK/TCM associated with the Ungerboeck 2-state 8-PSK/TCM (spectral efficiency 2 bits/s/Hz). Henceforth, a signal $s(i)$ will also be denoted by i . Indeed, $i'(k) = i(k) - i(k-1) \bmod 8$. The design of the differential TCM can be achieved by splitting each node in the trellis of the associate (conventional) TCM code (Fig.4). Unfortunately, we shall see that these codes are based on nonlinear convolutional codes. Despite this fact, decoding can be performed using the Viterbi algorithm.

Suppose, for instance, a code sequence $\{s_k\} = s^0 s^2 s^3 s^3 s^6 s^5 s^4 s^6 s^3 \dots$. Then, it can be easily checked that the respective differentially-encoded sequence, $\{s_k\} = s^0 s^2 s^5 s^0 s^6 s^3 s^7 s^5 s^0 \dots$, belongs to the associated Ungerboeck trellis code.

TABLE I. State Transitions for a 4-state Convolutional Code associated with the 2-state Ungerboeck Code.

| old state | s_{k-1} | new state | s_k | s'_k | state' (node) |
|-----------|-----------|-----------|-------|--------|---------------|
| 1 | 0,4 | 1 | 0,4 | 0,4 | 11 \equiv 1 |
| 1 | 0,4 | 2 | 2,6 | 2,6 | 12 \equiv 3 |
| 1 | 1,5 | 1 | 0,4 | 3,7 | 11 \equiv 1 |
| 1 | 1,5 | 2 | 2,6 | 1,5 | 12 \equiv 3 |
| 2 | 2,6 | 1 | 1,5 | 3,7 | 21 \equiv 2 |
| 2 | 2,6 | 2 | 3,7 | 1,5 | 22 \equiv 4 |
| 2 | 3,7 | 1 | 1,5 | 2,6 | 21 \equiv 2 |
| 2 | 3,7 | 2 | 3,7 | 0,4 | 22 \equiv 4 |

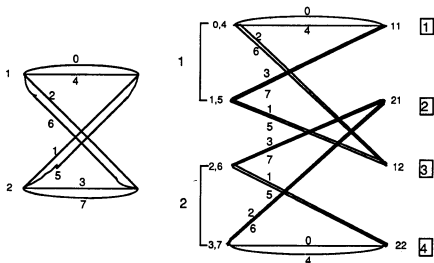


Figure 4. Trellis diagram for 2-state U code and its associated 4-state U' code.

TABLE II. State Transitions for the 4-state TCM/DPSK Code.

| state E ₁ E ₀ | input X ₂ X ₁ | output Y ₂ Y ₁ Y ₀ | new state E ₁ E ₀ |
|--|--|--|--|
| 00 1 | 00 | 000 (0) | 00 1 |
| 00 1 | 01 | 010 (2) | 10 3 |
| 00 1 | 10 | 100 (4) | 00 1 |
| 00 1 | 11 | 110 (6) | 10 3 |
| 01 2 | 00 | 001 (1) | 10 3 |
| 01 2 | 01 | 011 (3) | 00 1 |
| 01 2 | 10 | 101 (5) | 10 3 |
| 01 2 | 11 | 111 (7) | 00 1 |
| 10 3 | 00 | 011 (3) | 01 2 |
| 10 3 | 01 | 001 (1) | 11 4 |
| 10 3 | 10 | 111 (7) | 01 2 |
| 10 3 | 11 | 101 (5) | 11 4 |
| 11 4 | 00 | 000 (0) | 11 4 |
| 11 4 | 01 | 010 (2) | 01 2 |
| 11 4 | 10 | 100 (4) | 11 4 |
| 11 4 | 11 | 110 (6) | 01 2 |

We have used Karnaugh maps [KOR&KOR 1968] to minimize the Boolean functions: $Y_n^1 = f_1(X_1; E_1, E_0)$ and $Y_n^0 = f_0(X_1; E_1, E_0)$.

The corresponding truth tables are presented in figure 5. Finally, the minimal encoder of such a code is sketched in figure 6.

| Y_1 | | | Y_0 | | |
|--------------------------|---|---|--------------------------|---|---|
| $E_1 E_0 \backslash x_1$ | 0 | 1 | $E_1 E_0 \backslash x_1$ | 0 | 1 |
| 00 | 0 | 1 | 00 | 0 | 0 |
| 01 | 0 | 1 | 01 | 1 | 1 |
| 11 | 0 | 1 | 11 | 0 | 0 |
| 10 | 1 | 0 | 10 | 1 | 1 |

$$Y_1 = \bar{E}_1 X_1 + E_0 X_1 + E_1 \bar{E}_0 \bar{X}_1$$

$$= \bar{E}_1 E_0 \oplus X_1$$

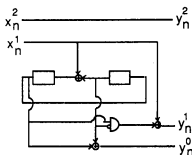
$$Y_0 = \bar{E}_1 E_0 + E_1 \bar{E}_0 = E_0 \oplus E_1$$
Figure 5. Truth tables for the output Boolean variables Y_n^1, Y_n^0 of the 4-state code.

Figure 6. 4-state nonlinear convolutional encoder for 8-DPSK/TCM.

• 8-state 8-DPSK/TCM

The approach to design an 8-state 8-DPSK/TCM follows in a quite similar way. We just outline the results. Tables III and IV present state transitions, figure 7 the trellis diagram, figure 8 truth tables and finally, the nonlinear convolutional encoder in shown in figure 9.

TABLE III. State Transitions for a 8-state Convolutional Code associated with the 4-state Ungerboeck Code.

| old state | s_{k-1} | new state | s_k | s'_k | state' (node) |
|-----------|-----------|-----------|-------|--------|---|
| 1 | 0,4 | 1 | 0,4 | 0,4 | 11 \equiv 1 |
| 1 | 0,4 | 2 | 2,6 | 2,6 | 12 \equiv 3 |
| 1 | 2,6 | 1 | 0,4 | 2,6 | 11 \equiv 1 |
| 1 | 2,6 | 2 | 2,6 | 0,4 | 12 \equiv 3 |
| 2 | 2,6 | 3 | 1,5 | 3,7 | 23 \equiv 5 |
| 2 | 2,6 | 4 | 3,7 | 1,5 | 24 \equiv 7 |
| 2 | 0,4 | 3 | 1,5 | 1,5 | 23 \equiv 5 |
| 2 | 0,4 | 4 | 3,7 | 3,7 | 24 \equiv 7 |
| 3 | 1,5 | 1 | 2,6 | 1,5 | 31 \equiv 2 |
| 3 | 1,5 | 2 | 0,4 | 3,7 | 32 \equiv 4 |
| 3 | 3,7 | 1 | 2,6 | 3,7 | 31 \equiv 2 |
| 3 | 3,7 | 2 | 0,4 | 1,5 | 32 \equiv 4 |
| 4 | 3,7 | 3 | 3,7 | 0,4 | 43 \equiv 6 |
| 4 | 3,7 | 4 | 1,5 | 2,6 | 44 \equiv 8 |
| 4 | 1,5 | 3 | 3,7 | 2,6 | 43 \equiv 6 |
| 4 | 1,5 | 4 | 1,5 | 0,4 | 44 \equiv 8 |

TABLE IV. State Transitions for the 8-state TCM/DPSK Code.

| state E ₁ E ₀ | | input X ₂ X ₁ | | output Y ₂ Y ₁ Y ₀ | | | new state E ₁ E ₀ | |
|--|---|--|--|--|-----|--|--|---|
| 000 | 1 | 00 | | 000 | (0) | | 000 | 1 |
| 000 | 1 | 01 | | 010 | (2) | | 010 | 3 |
| 000 | 1 | 10 | | 100 | (4) | | 000 | 1 |
| 000 | 1 | 11 | | 110 | (6) | | 010 | 3 |
| 001 | 2 | 00 | | 000 | (0) | | 010 | 3 |
| 001 | 2 | 01 | | 010 | (2) | | 000 | 1 |
| 001 | 2 | 10 | | 100 | (4) | | 010 | 3 |
| 001 | 2 | 11 | | 110 | (6) | | 000 | 1 |
| 010 | 3 | 00 | | 011 | (3) | | 100 | 5 |
| 010 | 3 | 01 | | 001 | (1) | | 110 | 7 |
| 010 | 3 | 10 | | 111 | (7) | | 100 | 5 |
| 010 | 3 | 11 | | 101 | (5) | | 110 | 7 |
| 011 | 4 | 00 | | 011 | (3) | | 110 | 7 |
| 011 | 4 | 01 | | 001 | (1) | | 100 | 5 |
| 011 | 4 | 10 | | 111 | (7) | | 110 | 7 |
| 011 | 4 | 11 | | 101 | (5) | | 100 | 5 |
| 100 | 5 | 00 | | 001 | (1) | | 001 | 2 |
| 100 | 5 | 01 | | 011 | (3) | | 011 | 4 |
| 100 | 5 | 10 | | 101 | (5) | | 001 | 2 |
| 100 | 5 | 11 | | 111 | (7) | | 011 | 4 |
| 101 | 6 | 00 | | 001 | (1) | | 011 | 4 |
| 101 | 6 | 01 | | 011 | (3) | | 001 | 2 |
| 101 | 6 | 10 | | 101 | (5) | | 011 | 4 |
| 101 | 6 | 11 | | 111 | (7) | | 001 | 2 |
| 110 | 7 | 00 | | 000 | (0) | | 101 | 6 |
| 110 | 7 | 01 | | 010 | (2) | | 111 | 8 |
| 110 | 7 | 10 | | 100 | (4) | | 101 | 6 |
| 110 | 7 | 11 | | 110 | (6) | | 111 | 8 |

| | | | |
|---|----|---------|---|
| 111 8 | 00 | 000 (0) | 111 8 |
| 111 8 | 01 | 010 (2) | 101 6 |
| 111 8 | 10 | 100 (4) | 111 8 |
| 111 8 | 11 | 110 (6) | 101 6 |

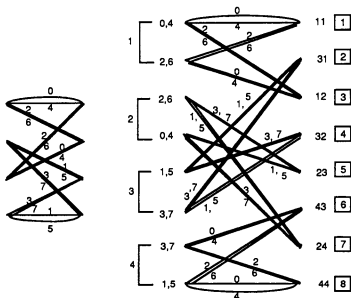


Figure 7. Trellis diagram for 4-state U code and its associated 8-state U' code.

| E_2E_1 | 00 | 01 | 11 | 10 |
|----------|----|----|----|----|
| E_0X_1 | 0 | 1 | 0 | 0 |
| 00 | 0 | 1 | 0 | 0 |
| 01 | 1 | 0 | 1 | 1 |
| 11 | 1 | 0 | 1 | 1 |
| 10 | 0 | 1 | 0 | 0 |

| E_2E_1 | 00 | 01 | 11 | 10 |
|----------|----|----|----|----|
| E_0X_1 | 0 | 1 | 0 | 1 |
| 00 | 0 | 1 | 0 | 1 |
| 01 | 0 | 1 | 0 | 1 |
| 11 | 0 | 1 | 0 | 1 |
| 10 | 0 | 1 | 0 | 1 |

Figure 8. Truth tables for the output Boolean variables Y_n^1 , Y_n^0 of the 8-state code.

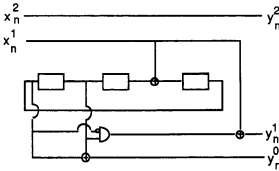


Figure 9. 8-state nonlinear convolutional encoder for 8-DPSK/TCM.

II.2.3. TRELLIS-CODED DIFFERENTIAL MODULATION.

The (noisy) received symbols are $z_k = s_k(i) e^{j\theta} + n_k$ $k=0,1,\dots,L-1$. Since we consider an AWGN channel, the optimal (maximum likelihood) decoding rule consists of minimizing the squared Euclidean distance between the received sequence and the phase shifted coded sequences, that is,

$$\begin{aligned} \text{MIN} \quad & \sum_{k=0}^{L-1} |z_k - s_k(i) e^{j\theta}|^2, \\ & \theta, i(k) \\ & k=0,1,\dots,L-1 \\ & \text{s.t.} \\ & \{s_k(i)\} \in \text{TCM}. \end{aligned} \quad (1)$$

where L is the length of the coded sequence corresponding to the decision delay.

In the case of multiphase signaling, the channel symbols have identical energy so that eq (1) reduces to

$$\begin{aligned} \text{MAX} \quad & \text{Re} \sum_{k=0}^{L-1} z_k s_k(i)^* e^{-j\theta}, \\ & \theta, \{i(k)\} \\ & \text{s.t.} \\ & \{s_k(i)\} \in \text{TCM}. \end{aligned} \quad (2)$$

where $*$ denotes the complex conjugate. For the sake of brevity, we drop the constraint on the maximization. Now applying the identity $\text{Re}(z_1 z_2) = \text{Re} z_1 \cdot \text{Re} z_2 - \text{Im} z_1 \cdot \text{Im} z_2$, we get

$$\begin{aligned} \text{MAX} \quad & \Re \left(\sum_{k=0}^{L-1} z_k s_k(i)^* \right) \cos \theta + \Im \left(\sum_{k=0}^{L-1} z_k s_k(i)^* \right) \sin \theta. \\ & \theta, \{ i(k) \} \end{aligned} \quad (3)$$

First, we solve this maximization on the phase by setting $\theta = \theta_{\text{opt}}$ such that the partial derivative regarding θ vanishes, i.e.,

$$-\Re \left(\sum_k z_k s_k(i)^* \right) \sin \theta_{\text{opt}} + \Im \left(\sum_k z_k s_k(i)^* \right) \cos \theta_{\text{opt}} = 0. \quad (4)$$

Therefore, we have

$$\theta_{\text{opt}} = \arctg \frac{\Im \left(\sum_k z_k s_k(i)^* \right)}{\Re \left(\sum_k z_k s_k(i)^* \right)}. \quad (5)$$

Using the polar representation of complex numbers, we found that

$$\theta_{\text{opt}} = \arg \left(\sum_k z_k s_k(i)^* \right), \quad (6)$$

which finally yields

$$\exp(j \theta_{\text{opt}}) = \frac{(\sum_k z_k s_k(i)^*)}{|\sum_k z_k s_k(i)^*|}. \quad (7)$$

Inserting the above phase estimation on eq(2), it follows that the optimal non-coherent decoding rule becomes

$$\begin{aligned} \text{MAX} \quad & \Re \left(\sum_k z_k s_k(i)^* \right) \frac{(\sum_m z_m^* s_m(i))}{|\sum_m z_m^* s_m(i)|}. \\ & \theta, \{ i(k) \} \end{aligned} \quad (8)$$

Consequently, the decoding is carried out by finding an information stream such that

$$\underset{\{i(k)\}}{\text{MAX}} \quad \left| \sum z_k s_k(i)^* \right|, \quad (9)$$

where the sequence $\{s_k(i)\}$ is the output of the TCM (a coded sequence) corresponding to an input sequence $\{i(k)\}$ (an uncoded sequence).

The problem here is that the metric (9) is not an additive one so that we cannot use the Viterbi algorithm. As a consequence, the optimal non-coherent decoding rule (exhaustive decoding) is impractical even for codes with short constraint lengths. In the following we present a simplification of such a decoding rule (a sub-optimum decoding) which gives a new interpretation of differential coding.

Let us consider the simplest phase estimation, namely phase differential estimation, for which we take into account only a single term in (7), according to

$$e^{j\theta_{\text{opt},k}} = \frac{(z_{k-1} s_{k-1}(i))^*}{|z_{k-1} s_{k-1}(i)|}. \quad (10)$$

This has the additional advantage of supposing a relatively constant carrier phase just over two consecutive signaling intervals, which is certainly a more realistic condition than to admit a stable phase over several signaling intervals. Furthermore, this could also be valuable over channels where the propagation delay changes with time.

Now using this approximation, the decoding rule (2) becomes

$$\text{MAX} \quad \Re e^{-j \sum_{k=0}^{L-1} z_k s_k(i)^* \frac{z_{k-1}^* s_{k-1}(i)}{|z_{k-1} s_{k-1}(i)|}}, \quad (11)$$

where the following definitions are used: $s_{-1}(i) = z_{-1}(i) := 1$.

It should be pointed out that we have now an additive metric, so the Viterbi decoding algorithm can be used.

We also recall the optimum coherent decoding rule (MLD) for trellis coded multiphase PSK modulation which is given by

$$\begin{aligned} \text{MAX} \quad & \text{Re} \sum z_k s_k(i)^* \\ \{s_k(i)\} \in & \text{TCM.} \end{aligned} \quad (12)$$

We suggest to define two new variables to be referred to as differential variables, according to

$$s'_k(i) := \frac{s_k(i) s_{k-1}(i)^*}{|s_{k-1}(i)|} \quad \text{and} \quad z'_k(i) := \frac{z_k(i) z_{k-1}(i)^*}{|z_{k-1}(i)|}. \quad (13)$$

Thus, the binary operation "prime" combines consecutive symbols and corresponds to a differential encoding. It can readily be noticed that $|s'_k(i)| = |s_k(i)|$ and $|z'_k(i)| = |z_k(i)|$. Figure 10 present schemes for both differential and inverse differential encoders.

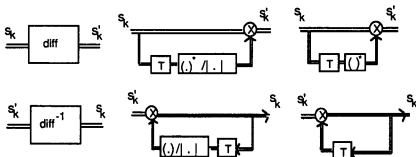


Figure 10. Schemes for direct and inverse differential encoders:

(a) black box representations, (b) general differential operation, (c) differential encoder for M-PSK.

The (sub-optimal) decoding rule (11) i.e., the differentially coherent decoding rule, can be rewritten as

$$\begin{aligned} \text{MAX} \quad & \text{Re} \sum z'_k s'_k(i)^* \\ \{s_k(i)\} \in & \text{TCM.} \end{aligned} \quad (14)$$

which has exactly the same mathematical expression as the coherent one (12) except the fact it involves differential symbols. More generally, even when channel symbols have

not the same energy, the use of the differentially approximated phase estimate (10) in the optimal non-coherent decoding rule (1) yields:

$$\text{MIN} \quad \sum_{k=0}^{L-1} |z'_k - s'_k(i)|^2. \quad (15)$$

Accordingly, we should use a trellis code which has minimal squared free Euclidean distance between sequences $\{s'_k(i_1)\}$ and $\{s'_k(i_2)\}$ as large as possible. To this end, an Ungerboeck code must be used. The communication system of Fig.1 has been simulated considering the TCM as an Ungerboeck code, and the resulting error performance has been evaluated. The one-sigma statistical confidence interval as a fraction of the sample mean has been kept within 10%.

As a result, we found that Ungerboeck codes are "invariant" to differential modulation in the sense that they are yet very good codes. This fact is also confirmed by the computer simulations (see Fig.11). It can be seen that the gain of 8-DPSK/TCM with respect to the theoretical uncoded 4-DPSK is practically the same as the coding gain of 8-PSK/TCM with respect to the uncoded 4-PSK. Furthermore, the schemes used have small complexity in comparison with other similar schemes [EDB 1987a, EDB 1987b].

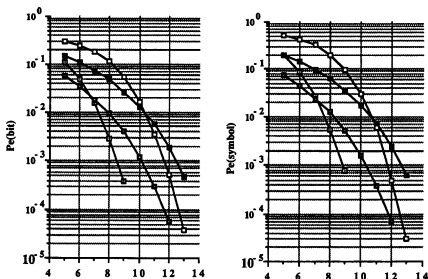


Figure 11. Simulated error rate vs signal-to-noise ratio (dB). Points correspond to uncoded 4-DPSK (shaded square); 4-state TCM / 8-DPSK (white square); uncoded

4-PSK (almost shaded square), and 4-state TCM / 8-PSK (almost white square).

II.2.4. CONCLUSIONS.

We have presented some aspects of the noncoherent detection of TCM coded systems. In particular, the design of TCM codes for DPSK modulation using differential coherent detection over the Gaussian channel has been considered. A family of nonlinear TCM schemes having twice the number of states of the associated Ungerboeck code has been introduced. By interpreting differential coherent detection as a first order approximation of the optimal noncoherent detection, we have found that Ungerboeck codes are still the best ones. A simple 4-state TCM / 8-PSK Ungerboeck code applied to DPSK modulation with differential coherent detection provides essentially the same performance as the uncoded coherent 4-PSK system at 10^{-5} . A 4-state trellis complexity suffices to avoid the need of a coherent phase reference. Gains achieved with respect to an uncoded system having the same spectral efficiency and no requirements of coherent phase (e.g., 4-DPSK) are given.

REFERENCES:

- [DIV&SIM 1988a], [DIV&SIM 1988b], [SIM&DIV 1987], [UNG 1982], [UNG 1987a,1987b], [JAM&Le_N 1991], [LEE&FAR 1991], [MAK&FER 1988], [EDB 1987a, EDB 1987b], [McL *et al.* 1988].

CHAPITRE III

A PROPOS DE L'EXISTENCE DE BONS CODES DE RESEAU

III.1. A CAPACITY THEOREM FOR LATTICE CODES OVER THE GAUSSIAN CHANNEL

"On se persuade mieux, pour l'ordinaire, par des raisons
qu'on a soi même trouvées, que par celles qui sont
venues dans l'esprit des autres" B. PASCAL.

Il y a déjà longtemps que les réseaux, au sens mathématique et cristallographique du terme, ont été proposés comme technique de codage sur le canal gaussien [e.g. BLA 1971, LEECH&SLO 1971]. Cependant, c'est seulement après la récente publication d'un article de Calderbank et Sloane [CAL&SLO 1987] sur les codes de réseau et de classes latérales qu'ils ont suscité un grand intérêt. Il y a maintenant consensus sur le fait que ces codes constituent une technique puissante dans le domaine de la modulation codée. Son application pratique a été immédiate puisque, dès la fin des années 80, les réseaux de Gosset et de Leech ont été utilisés dans un MODEM numérique (chez Motorola) sur le canal téléphonique [LAN&LON 1989].

L'intérêt pour ce type de codage remonte en particulier à une étude de R. de Buda [deB 1975], qui a montré en appliquant le théorème de géométrie des nombres de Minkowski-Hlawka [CAS 1959] qu'il existe des codes de réseau dont les performances sont très proches de la limite théorique (la capacité du canal).

J'ai repris et modifié ces résultats pour établir un théorème sur la capacité du canal gaussien quand des réseaux de nombre de dimensions élevé sont utilisés pour définir les constellations.

J'ai repris le raisonnement sur le "durcissement" des hypersphères quand le nombre de dimensions tend vers infini. J'ai démontré que cela résulte de la loi faible des grands nombres pour des variables aléatoires de distribution en χ^2 , en appliquant la théorie des distributions [BOU 1964].

D'autre part, j'ai repris l'interprétation géométrique bien connue d'entassement de sphères [SHA 1949] pour l'adapter au cas des codes de réseau. Cette approche a permis de démontrer un théorème où est explicité le rapport entre le paramètre a du réseau (rayon de la sphère de même volume que la région de Voronoi) et le débit R en bits par dimension. Ce théorème aboutit à deux nouvelles bornes sur la probabilité d'erreur des codes de réseau, qui jouent un rôle essentiel dans la démonstration du théorème de capacité.

J'ai ensuite démontré un théorème de capacité pour les codes de réseau sur un canal à bruit additif, gaussien et blanc: d'abord, l'assertion négative (*strong converse*) que la probabilité d'erreur tend vers 1 quand le débit est supérieur à la capacité, quel que soit le réseau utilisé; ensuite, l'assertion positive (*direct statement*) de l'existence d'un réseau dont la probabilité d'erreur est aussi petite qu'on le souhaite quand le débit est inférieur à la capacité. Tout comme dans le théorème de Shannon, ces résultats montrent l'existence de bons codes mais ne donnent pas de solution explicite du problème. Cependant, si le codage aléatoire n'est pas vraiment réalisable car l'absence de structure, empêchant toute forme de simplification, rend le décodage beaucoup trop complexe pour une réalisation effective, les codes de réseau ont au contraire une forte structure algébrique et géométrique qui peut être utilisée pour simplifier la tâche de décodage. Cela s'est bien vérifié dans le cas du réseau de Leech pour lequel de nouveaux algorithmes rapides viennent d'être introduits et permettent

une diminution étonnante de la complexité du décodeur [cf. I.4.2.1]. Finalement, j'ai obtenu une expression explicite d'une fonction de fiabilité [VIT&OMU 191982, GAL 1968] $E(R,C)$ pour les codes de réseau sur un canal gaussien en montrant qu'il existe des codes dont la probabilité d'erreur est bornée exponentiellement par

$$P_e < \exp(-nE(R,C)),$$

où n est la dimension du réseau et C la capacité du canal.

ABSTRACT

A capacity theorem for lattice codes signaling is presented which is based on an upper bound on the error probability introduced by R. de Buda. It is shown that lattice codes can be used to achieve the channel capacity for any signal-to-noise ratio (direct statement) and the negative statement of the capacity theorem is also proved (a strong converse to the capacity theorem). The sphere hardening is shown to result from the weak law of large numbers. The proof allows a better understanding of the application of dense lattices as an efficient signaling alphabet. An expression of the reliability function $E(R,C)$ for lattices in AWGN channels is also presented.

III.1.1. INTRODUCTION

A LONG TIME PASSED since Shannon's work [1] before the transmission at rates close to the channel capacity (near the cut-off rate) becomes a reality. The conventional coding techniques introduced during three decades were shown to be useful only for channels in the power-limited region. It was thought that coding had very little to offer to channels in the band-limited region [2]. A breakpoint was finally introduced in Ungerboeck's famous paper on set-partition coding [3]. A new research area in coding theory namely, coded-modulation, was created which has met tremendous interest and immediate practical applications, specially in digital MODEMs [4].

Although dense lattices have long been suggested as efficient codes in Gaussian channels [5,6,7] it seems they became attractive after an excellent paper by Calderbank and Sloane [8]. Since then, coding theorists had increased attention to lattices and cosets. These are an alternative way to represent the coded modulation systems: coset codes (lattice codes) are a powerful and attractive representation of trellis-coded (block-coded) modulation. The main fundamentals on cosets and lattices as a coding technique were recently introduced by Forney [9,10]. The interest in lattice codes stemmed from the work of de Buda [11] who showed that high dimensional lattices can be used in order to reduce the error probability in the presence of Gaussian noise and that such a code asymptotically achieves the channel capacity at high signal-to-noise ratio. In the present paper, we show that lattice codes can achieve the capacity at any signal-to-noise ratio and we also prove the negative statement (strong converse) of the capacity theorem. These results confirm those discussed by de Buda in a more recent paper

[12], but in a simpler way which moreover relies on the same tools as he used in his earlier work.

The paper is organized as follows. Section III.1.2 introduces upper and lower bounds on the worst case error probability which are a slight modification of those presented by de Buda [11]. Sphere hardening theorems are presented in section III.1.3 and they are used, in section III.1.4, to prove a capacity theorem for lattice codes over the Gaussian channel. Finally, a reliability function for AWGN channels is introduced in section III.1.5 assuming that a lattice code is used.

III.1.2. BOUNDS ON THE ERROR PROBABILITY

Some bounds introduced here are essentially a modified version of those presented by R. de Buda. Let $J_n(r)$ denote the volume of a hypersphere of center \underline{x} and radius r , $B_n(\underline{x}, r) = \{\underline{x}' \in \mathbb{R}^n \mid |\underline{x}' - \underline{x}| \leq r\}$, which is given by

$$J_n(r) = V_n r^n \quad (1)$$

where $V_n = \pi^{n/2} / \Gamma(\frac{n}{2} + 1)$ is the volume of a normalized n -sphere.

Let us suppose now that the n -dimensional bandlimited noise is AWGN with average power N , so that

$$\underline{r} = (n_1, n_2, \dots, n_n) \quad (2)$$

where $\{n_i\}$ are independent identically distributed (i.i.d.) Gaussian random variables of zero mean and variance N_n (denoted by $n_i \sim N(0, N_n)$). Here, N_n denotes the variance (average power) of the noise in each dimension. Of course, since the n -dimensional noise is isotropic, one has $N_n = N/n$.

Let r be the magnitude of the noise vector. Then

$$r^2 = n_1^2 + n_2^2 + \dots + n_n^2 = |\underline{r}|^2. \quad (3)$$

The n -variate probability density of \underline{r} is given by

$$P_n(\underline{r}) = (2\pi N_n)^{-n/2} \exp(-r^2/2N_n) \quad (4)$$

and the corresponding normalized probability density is

$$Z_n(\underline{r}) \equiv (2\pi)^{-n/2} \exp(-r^2/2). \quad (5)$$

The n -variate chi-square distribution $Q(\chi^2|n)$ is the probability that the sum of n squared normalized Gaussian random variables exceeds χ^2 . It is convenient to express $Q(\chi^2|n)$ as:

$$Q(\chi^2|n) = \int_{\underline{r}}^{\infty} Z_n(\underline{r}) dJ_n(\underline{r}). \quad (6)$$

It is well known that maximum likelihood decoding (MLD) in AWGN is equivalent to choose the signal point closest to the received signal. Thus, the decision region of each

lattice point \underline{x}_k coincides with its Voronoi region $V(\underline{x}_k)$. Given that a lattice point $\underline{x} = \underline{x}_k$ was transmitted, the conditional error probability is

$$P_E(\underline{x}) = P(\text{outside } V(\underline{x}_k) | \underline{x}_k) = 1 - \int_{V(\underline{x}_k)} P_n(\underline{x}_k + \underline{r}) dV. \quad (7)$$

In an unbounded lattice, because of both the lattice symmetry and the additiveness of the noise, it follows that

$$P_E(\underline{x}) = 1 - \int_{V(0)} P_n(\underline{r}) dV = P_E(0). \quad (8)$$

Then the probability that the noise alone is outside the Voronoi region surrounding the origin is the error probability P_E of any lattice point, that is, $(\forall \underline{x} \in \Lambda) \quad P_E(\underline{x}) = P_E$. If the signaling set Ω (coded constellation) is given by all the points of a lattice contained in a bounded region, the outmost signal points have Voronoi regions which are larger than $V(0)$ so they have an error probability less than P_E . Thus, P_E is the worst case error probability. It can be bounded as follows.

Lower Bound

Let $\{\underline{x}_k\}$ be points of a lattice Λ , and let us consider n -dimensional spheres of radius a , $A(\underline{x}_k) = B_n(\underline{x}_k, a)$, having the same volume as a Voronoi region namely, $V(\underline{x}_k) = V(0) = \det \Lambda$, that is, such that

$$J_n(a) = \det \Lambda. \quad (9)$$

Since $P_n(\underline{r})$ is a monotonically decreasing function of r , it follows that

$$P(\text{inside } A(0) | 0) = \int_{A(0)} P_n(\underline{r}) dV \geq P(\text{inside } V(0) | 0) = \int_{V(0)} P_n(\underline{r}) dV. \quad (10)$$

A simple bound on P_E can be obtained by taking the complementary event in R^n , i.e., $P(\text{outside } A(0) | 0) \leq P(\text{outside } V(0) | 0)$, so that

$$\int_a^\infty P_n(r) dJ_n(r) \leq P_E \quad (11)$$

or, after changing the integration variable

$$P_E \geq \int_{a/\sqrt{N_n}}^\infty Z_n(r') dJ_n(r') = Q(a^2/N_n | n). \quad (12)$$

Upper Bounds

A simple upper bound on P_E can be obtained by taking a hypersphere $B(0) = B_n(0, \rho)$ with center at the origin and radius $\rho = d_{\min}/2$, where d_{\min} is the minimum distance between the lattice points. Let $B^c(0)$ be the complementary set of such a hypersphere. Obviously, $P(\text{outside } V(0) | 0) \leq P(\text{outside } B(0) | 0)$ or

$$P_E \leq \int_{B^*(0)} P_n(\underline{r}) dV = \int_{\rho}^{\infty} P_n(r) dJ_n(r) . \quad (13)$$

Finally, we obtain

$$P_E \leq \int_{\rho/\sqrt{N_n}}^{\infty} Z_n(r') dJ_n(r') = Q(\rho^2/N_n |n) . \quad (14)$$

From (9) and the definition of the lattice density Δ [13], the relationship between ρ^2 and a^2 can easily be shown to be $\rho^2 = \Delta^{2/n} a^2$, so (14) may be rewritten as

$$P_E \leq Q(\Delta^{2/n} a^2/N_n |n) . \quad (15)$$

At this point, it is worthwhile to mention that (14) may be interpreted in a very interesting way. The signaling set Ω (coded constellation) is given by all points in a bounded lattice with average power $P(\Omega)$. Then

$$P_E \leq Q\left(\frac{\rho^2}{P(\Omega)} \frac{P(\Omega)}{N} n |n\right) = Q\left(\frac{\rho^2}{P(\Omega)} \gamma_{av} n |n\right) . \quad (16)$$

A figure of merit of the coded constellation is given by $FM(\Omega)$, where

$$FM(\Omega) =: \frac{\rho^2}{P(\Omega)} = \frac{(d_{\min}^2/2)}{P(\Omega)} . \quad (17)$$

Let us now suppose that decoding is effected by taking hyperspheres $B(\underline{x}_k) = B_n(\underline{x}_k, b)$ with center at the lattice points \underline{x}_k as decision regions. The radius $r = b < a$ of these spheres will be selected later. Two types of error can occur with probabilities P_{EI} and P_{EII} , respectively. First, when the received signal is outside the correct decision sphere. Secondly, since the decisions sphere overlap, when a signal from a wrong lattice point is found inside the decision sphere, such an ambiguous decision will be counted as an error of type II.

The type I error probability P_{EI} can easily be determined by

$$P_{EI} = \int_{\rho}^{\infty} P_n(r) dJ_n(r) = Q(r^2/N_n |n) . \quad (18)$$

The type II error probability is bounded in a more complicated way, according to de Buda's bound. Let $f: \Lambda \rightarrow \mathbb{R}$ be a Riemann integrable function and let us denote by

$$f(\Lambda) = \sum_{\substack{\underline{x} \in \Lambda \\ \underline{x} \neq 0}} f(\underline{x}) \quad (19)$$

the sum of $f(\underline{x})$ over all lattice points except the origin. For a lattice point $\underline{x} \neq 0$, the type II error probability is

$$P_{EII}(\underline{x}) = \int_{|\underline{r}| < r} P_n(\underline{x} - \underline{r}) dV(\underline{r}) = f(\underline{x}) \quad (20)$$

so that

$$P_E II(\underline{x}) = f(\Lambda) . \quad (21)$$

In order to obtain a bound over $P_E II$, it is intended to use the following theorem of the geometry of numbers [11]:

THEOREM 1 (Minkowsky-Hlawka/de Buda): With $f(\Lambda)$ defined as above, given $\epsilon > 0$, there exists a lattice with $\det \Lambda = \epsilon$ such that $\det \Lambda \cdot f(\Lambda) \leq \int_{(\underline{x})} f(\underline{x}) dV(\underline{x})$

The first step for determining the bound on $P_E II$ is the evaluation of the integral

$$I = \int_{(\underline{x})} f(\underline{x}) dV(\underline{x}) \quad (22)$$

where f is given by (20). Substituting (20) in (22) and interchanging the order of the integrations (which is possible due the steep decrease of $P_n(\cdot)$), we have

$$I = \int_{|r| < r} \int_{(\underline{x})} P_n(\underline{x}-r) dV(\underline{x}) dV(r) . \quad (23)$$

Since $P_n(\cdot)$ is a probability density, the inner integral is unity, so

$$I = \int_{|r| < r} dV(r) = J_n(r) . \quad (24)$$

The application of (22),(24) and Theorem 1 implies that a lattice exists which verifies

$$f(\Lambda) \cdot J_n(a) \leq J_n(r) . \quad (25)$$

We can now apply the union bound and (21) so P_E is upper bounded by

$$P_E \leq P_E I + P_E II = Q(r^2/N_n |n) + J_n(r)/J_n(a) . \quad (26)$$

The parameter r (radius of a decision region) will be given the value b such that the tightest bound is obtained. It is a solution of

$$\frac{d(P_E I + P_E II)}{dJ_n(r/\sqrt{N_n})} = -Z_n(r/\sqrt{N_n}) + 1/J_n(a/\sqrt{N_n}) = 0 . \quad (27)$$

This equation has a single real positive root $r = b$, namely:

$$b = \sqrt{N_n(2 \ln J_n(a/\sqrt{N_n}) - n \ln(2\pi))} . \quad (28)$$

The tightest upper bound is thus

$$P_E \leq Q(b^2/N_n |n) + Z_n(b/\sqrt{N_n}) \cdot J_n(b) = Q(b^2/N_n |n+2) \quad (29)$$

where equality results from integration by part (see also Abramowitz-Stegun [14]). Substituting the volume of a normalized n -sphere and rearranging, we obtain

$$b^2/N_n = n \ln \left(\frac{n/2}{\Gamma^{2/n}(\frac{n}{2}+1)} \frac{a^2}{N} \right) . \quad (30)$$

III.1.3. SPHERE HARDENING THEOREMS

We are now interested in the behaviour of $Q(\chi^2|n)$ for n very large, where $\chi^2 = r^2$ is given by (3). Let us consider the following random variable transformations:

TABLE I. RANDOM VARIABLE TRANSFORMATIONS.

| r.v. | distribution | probability density |
|---------|--------------|--|
| r | χ_n | $P_r(\xi;n) = \frac{\xi^{n-1} \exp(-\xi^2/2)}{2^{\frac{n}{2}-1} \Gamma(\frac{n}{2})}$ |
| r^2 | χ_n^2 | $P_{r^2}(\xi;n) = \frac{\xi^{\frac{n}{2}-1} \exp(-\xi/2)}{2^{n/2} \Gamma(\frac{n}{2})}$ |
| r^2/n | χ_n^2/n | $P_{r^2/n}(\xi;n) = \frac{(n/2)^{n/2}}{\Gamma(n/2)} \xi^{\frac{n}{2}-1} \exp(-\frac{n}{2}\xi)$ |

where

$$P_{r^2/n}(\xi;n) = P_r(\sqrt{n\xi}; n) \frac{\sqrt{n}}{2\sqrt{\xi}} = P_{r^2}(n\xi; n) \cdot n. \quad (31)$$

Thus, $Q(\chi^2|n)$ may be expressed alternatively by:

$$Q(\chi^2|n) = \int_{\chi^2}^{\infty} P_{r^2}(\xi;n) d\xi = \int_{\chi^2/n}^{\infty} P_{r^2/n}(\xi;n) d\xi. \quad (32)$$

THEOREM 2: For any real λ^2 , the chi-square function verifies

$$\lim_{n \rightarrow \infty} Q(\lambda^2 n | n) = I_{(-\infty, 1)}(\lambda^2) =: \begin{cases} 0 & \text{if } \lambda^2 > 1, \\ 1 & \text{otherwise.} \end{cases}$$

The proof is based on a few results which allow a better understanding of the sphere hardening phenomenon and will be discussed later.

THEOREM 3 (Weak law theorem): Let $\{X_i\}$ be i.i.d. random variables with $X_i \sim N(0, \sigma^2)$ and $W = 1/n \sum_{i=1}^n X_i^2$. Then the limit in probability of W is σ^2 .

proof. The mean and variance of such a distribution can be shown to be $E(W) = \sigma^2$ and $\text{Var}(W) = 2(\sigma^2)^2/n$. By the Chebyshev inequality it follows that $P(|W - \sigma^2| > \varepsilon) \leq 2(\sigma^2)^2/(n\varepsilon^2) \rightarrow 0$ as $n \rightarrow \infty$. **Q.E.D. ■**

The Theorem above can be used to show that $P_{r^2/n}(\xi; n)$ defines a generalized function as we see by the following

THEOREM 4: The sequence of functions $\{P_{r^2/n}(\xi; n)\}$ defines the Dirac distribution $\delta(\xi-1)$.

proof. We know that

$$\int_{-\infty}^{\infty} P_{r^2/n}(\xi; n) d\xi = 1 \quad (33)$$

because $P_{r^2/n}(\cdot; n)$ is a pdf. On the other hand, from Theorem 3:

$$1 \geq \int_{1-\epsilon}^{1+\epsilon} P_{r^2/n}(\xi; n) d\xi = 1 - P(|W-1| > \epsilon) \geq 1 - 2/(n\epsilon^2).$$

Taking the limit (when n increases indefinitely) in the expression above,

$$\lim_{n \rightarrow \infty} \int_{1-\epsilon}^{1+\epsilon} P_{r^2/n}(\xi; n) d\xi = 1. \quad (34)$$

Then $\lim_{n \rightarrow \infty} P_{r^2/n}(\xi; n) = \delta(\xi-1)$ follows from (33) and (34). **Q.E.D. ■**

proof of Theorem 2:

Writing $Q(\lambda^2 n | n)$ in terms of the density (31) of the r.v. r^2/n , we have

$$\begin{aligned} Q(\lambda^2 n | n) &= \int_{\lambda^2}^{\infty} P_{r^2/n}(\xi; n) d\xi \quad \text{or} \\ \lim_{n \rightarrow \infty} Q(\lambda^2 n | n) &= \lim_{n \rightarrow \infty} \int_{\lambda^2}^{\infty} P_{r^2/n}(\xi; n) d\xi. \end{aligned} \quad (35)$$

The application of Theorem 4 yields:

$$\lim_{n \rightarrow \infty} Q(\lambda^2 n | n) = \int_{\lambda^2}^{\infty} \delta(\xi-1) d\xi \text{ and the result follows. } \quad \text{Q.E.D. } \blacksquare$$

LEMMA 5 : For any integer v , $\lim_{n \rightarrow \infty} Q(\lambda^2 n | n+v) = I_{(-\infty, 1)}(\lambda^2)$

proof. Once again, we write $Q(\chi^2 | n)$ in terms of $P_{r^2/n}(\cdot; n)$

$$\lim_{n \rightarrow \infty} Q(\lambda^2 n | n+v) = \lim_{n \rightarrow \infty} \int_{\lambda^2}^{\infty} P_{r^2/n}(\xi; n+v) d\xi \quad (36)$$

$$\text{but } P_{r^2/n}(\xi; n+v) = P_{r^2}(n\xi; n+v) \cdot n. \quad (37)$$

When n grows, $P_{r^2}(\cdot; n) \rightarrow P_{r^2}(\cdot; n+v)$, so that $\lim_{n \rightarrow \infty} P_{r^2/n}(\cdot; n+v) = \lim_{n \rightarrow \infty} P_{r^2/n}(\cdot; n)$ and the proof follows from Theorem 2. **Q.E.D. ■**

III.1.4. CODING THEOREMS FOR LATTICE CODES

The bounds introduced in section II depend on two parameters a and b , so we must find the relationship between these parameters and the transmission rate R . This can be done with the aid of a geometrical representation of signals and noise as shown in Figure 2.

The volume of a Voronoi region is $V(0) = \det \Lambda$ and we took spheres with same volume as $V(0)$ according to (9). If the transmission rate is R bits/dimension, then the total number of such n -spheres (or signal points) is 2^{nR} . The total volume occupied by all the Voronoi regions is $2^{nR} \cdot \det \Lambda = 2^{nR} \cdot J_n(a)$. Suppose that the code consists of all lattice points inside a hypersphere of radius $\sqrt{a^2 + S}$ (see Fig. 2), so the available volume is $J_n(\sqrt{a^2 + S})$. Then $V_n(\sqrt{a^2 + S})^n = 2^{nR} V_n a^n$, so that

$$R = \frac{1}{2} \log_2((a^2 + S)/a^2). \quad (38)$$

THEOREM 6: Given a lattice Λ , if the rate R is less than the channel capacity C then $a^2/N \geq 2^{2(C-R)}$. In contrast, if the rate R exceeds the capacity C , then $a^2/N \leq 2^{2(C-R)}$.

proof. The relationship between the lattice parameter a and the transmission rate R is given by (38) which implies $a^2 = S(2^{2R} - 1)^{-1}$. On the other hand, Shannon's capacity formula leads to $S/N = (2^{2C} - 1)$. Thus,

$$a^2/N = (2^{2C} - 1)(2^{2R} - 1)^{-1}. \quad (39)$$

This relation can be rewritten as

$$\frac{a^2}{N} = \frac{2^{2C} - 1}{2^{2C} 2^{2(R-C)} - 1} = 2^{2(R-C)} \frac{2^{2C} - 1}{2^{2C} - 2^{2(C-R)}}. \quad (40)$$

The proof is completed by comparing the second term of the right side of (40) with unity. **Q.E.D. ■**

Recollecting the bounds obtained in section II,

$$Q(a^2/N_n | n) \leq P_E \leq Q(b^2/N_n | n+2). \quad (41)$$

The proof of the capacity theorem is based on new bounds derived from these ones and theorem 6 as follows.

THEOREM 7 (Capacity Theorem)

Positive statement: If the rate R in bits/dimension is less than C , some lattice code exists such that the decoder error probability tends to zero as n grows.

Negative statement: Conversely, if R exceeds C , any lattice code has error probability near one.

proof.

(negative Statement) Any lattice code with $R > C$ verifies the inequality $a^2/N n \leq 2^{2(C-R)n} =: \alpha^2/N n$ which is an immediate consequence of Theorem 6. Then the worst case error probability for a lattice code obeys the following lower bounds:

$$P_E \geq Q(a^2/N_n | n) \geq Q(\alpha^2/N_n | n). \quad (42)$$

By hypothesis, $R > C$, so that $R = C + \varepsilon/2$ for some $\varepsilon > 0$. Applying the lower bound (42), any lattice code must verify $P_E \geq Q(2^{2(C-R)n} | n) = Q(2^{-\varepsilon n} | n)$. When the dimensionality increases without limit, then $\lim_{n \rightarrow \infty} P_E \geq \lim_{n \rightarrow \infty} Q(2^{-\varepsilon n} | n) = 1$ (from theorem 2).

(Positive statement) Analogically, the application of Theorem 6 to the lattice code with rate $R < C$ yields the following inequality:

$$b^2/N_n \geq 2n \ln 2 \left[C - R - \frac{1}{2} \log_2 \frac{\Gamma^{2/n}(\frac{n}{2}+1)}{n/2} \right] =: \beta^2/N_n \quad (43)$$

so that the worst case error probability verifies the upper bounds: $P_E \leq Q(b^2/N_n | n+2) \leq Q(\beta^2/N_n | n+2)$.

Let us now examine the condition $\beta^2/N > 1$. From (43) this implies

$$C - \frac{1}{2} \log_2 \frac{e \Gamma^{2/n}(\frac{n}{2}+1)}{n/2} - R > 0. \quad (44)$$

Let us define

$$C(n) =: C - \frac{1}{2} \log_2 \frac{e \Gamma^{2/n}(\frac{n}{2}+1)}{n/2}. \quad (45)$$

Then $C(n) = C + \phi(n)$, where $\phi(n)$ is a function which approaches zero when n increases indefinitely, as we shall see by

LEMMA 8 : $C(n)$ approaches the channel capacity C when $n \rightarrow \infty$, i.e., $\lim_{n \rightarrow \infty} C(n) = C$ bits/dimension.

proof. By the Stirling approximation, $\Gamma(\frac{n}{2}+1) \approx \sqrt{\pi n} (n/2e)^{n/2}$ so that

$$\frac{e \Gamma^{2/n}(\frac{n}{2}+1)}{n/2} \approx (\pi n)^{1/n}, \quad n \text{ large}.$$

Therefore, $\phi(n) \approx \frac{1}{n} \log_2(\pi n)$ goes to zero as $n \rightarrow \infty$, and the proof follows directly from the definition of $C(n)$. **Q.E.D. ■**

Rewriting the upper bound on P_E ,

$$P_E \leq Q(\beta^2/N \mid n+2) = Q((2\ln 2(C(n)-R)+1) \mid n+2). \quad (46)$$

The rate is less than C , so $R = C - \epsilon$ for some $\epsilon > 0$. From Theorem 1, a lattice code exists such that:

$$\beta^2/N = 2 \ln 2 (\epsilon - \phi(n)) + 1 \quad \text{and} \quad \lim_{n \rightarrow \infty} \beta^2/N = 2 \epsilon \ln 2 + 1.$$

Applying the upper bound (46) when $n \rightarrow \infty$ gives

$$\lim_{n \rightarrow \infty} P_E \leq \lim_{n \rightarrow \infty} Q(\beta^2/N_n \mid n+2) = \lim_{n \rightarrow \infty} Q[(1+2\epsilon \ln 2) \mid n+2] \rightarrow 0$$

(from Lemma 5) **Q.E.D. ■**

This result has a well known geometrical interpretation. If $a^2 \rightarrow N$, then $R \rightarrow C$ follows from (38). Furthermore, $\beta^2 \rightarrow b^2$ and $\alpha^2 \rightarrow a^2$ follows from (40) and (43). If $a^2 \rightarrow N$, then $b \rightarrow a$ when $n \rightarrow \infty$ and the upper and lower bounds in (41) agree. For high dimensionality, the noise becomes concentrated on the surface of a hypersphere of radius equal to the average power of the noise (consequence of Theorem 4). Each lattice point is the center of a hypersphere of radius b ($b \rightarrow a$). If the signal plus noise is in the n -sphere around a lattice point \underline{x} , it will be decoded as \underline{x} . Then if $a^2 > N$ ($b^2 \rightarrow a^2 > N$), the received signal will be at the surface of the hypersphere around the transmitted lattice point as $n \rightarrow \infty$ and no decoding errors will occur. In contrast, if $a^2 < N$, then $b^2 \rightarrow a^2 < N$ and the signal plus noise points will be almost surely decoded in a wrong way. To summarize, if $a^2 > N$ then $P_E \rightarrow 0$ as n grows, otherwise $P_E \rightarrow 1$.

The use of dense lattice codes is explained as follows. Let us suppose that the n -dimensional lattice was obtained by rescaling [13] a lattice Λ^* , i.e., $\Lambda = (\rho/\rho^*)\Lambda^*$.

We can rewrite equation (9) as $a^2 = (\det \Lambda / V_n)^{2/n}$ so that $a^2 > N$ implies

$$(\det \Lambda / V_n)^{2/n} > N. \quad (47)$$

But $\det \Lambda = (\rho/\rho^*)^n \det \Lambda^*$, hence the above equation yields

$$\rho^2 (\Delta^*)^{-2/n} > N. \quad (48)$$

Then we must take a compact lattice Λ^* and use a rescaled version Λ so that the minimal distance 2ρ between lattice points obeys (48)

III.1.5. PERFORMANCE OF LATTICE CODES OVER THE AWGN CHANNEL

A reliability function for the AWGN channel can also be obtained by

THEOREM 9 : If $R < C$ bits/dim, then a lattice code exists such that

$$P_E < \frac{1}{2} \exp(-nE(R, C))$$

as $n \rightarrow \infty$, where $E(R, C) = E(R - C) = (\sqrt{1+2 \ln 2 (C-R)} - 1)^2$.

proof. For $\chi^2 > n + 2$, n large, the following approximation of the chi-square distribution can be used:

$$Q(\chi^2 | n+2) \approx \frac{1}{2} \operatorname{erfc}(\sqrt{\chi^2 - \sqrt{n+1.5}}) < \frac{1}{2} \exp(-(\sqrt{\chi^2 - \sqrt{n+1.5}})^2).$$

Let $\chi^2 = (2\ln 2 \cdot (C(n) - R) + 1)n$. We find it convenient to express χ^2 as

$$\chi^2 = (1 + 2 \ln 2 \cdot (C - R - \phi(n) - \frac{1}{n} \log_2 e))n + 2.$$

If $R < C$, there is $\epsilon > 0$ such that $C - R = \epsilon$. For a lattice with a large enough dimension n , $\epsilon > \phi(n) + \frac{1}{n} \log_2 e$, so that $\chi^2 > n + 2$. In this case, the chi-square approximation gives

$$P_E \leq \frac{1}{2} \exp[-n(\sqrt{2\ln 2 \cdot (C(n) - R) + 1} - \sqrt{1 + 1.5/n})^2]. \quad (49)$$

Let us consider the function $E_n(R, C)$ defined as

$$E_n(R, C) = (\sqrt{2\ln 2 \cdot (C(n) - R) + 1} - \sqrt{1 + 1.5/n})^2.$$

When n grows, this results in

$$\lim E_n(R, C) = (\sqrt{1 + 2\ln 2 \cdot (C - R)} - 1)^2 \text{ as } n \rightarrow \infty. \quad (50)$$

If we define the reliability function $E(R, C)$ as the above limit, it follows that a lattice exists for which the error probability is bounded as proposed. **Q.E.D. ■**

It is interesting to note that even for relatively small n , the approximation $E_n(C - R) \approx E(C - R)$ is remarkably tight. We present in Figure 3 the $E(R, C)$ function for an AWGN channel with SNR = 27.1 dB as an example. We have used as a bound $P_E < \frac{1}{2} 2^{-nE_2(R, C)}$ instead of $\frac{1}{2} \exp[-nE(R, C)]$, where $E_2(R, C) = E(R, C)/\ln(2)$.

It is now well established [15,16] that the Leech lattice can be used on telephone channels ($B = 2742.86$ Hz, SNR = 27 dB) for transmitting at $R = 3.5$ bits/dimension (or 7 (bits/s)/Hz) with an acceptable error probability, which yields a 19,200 bits/s digital MODEM. It can be seen that lattices are powerful coding tools to allow reliable transmission at high rates (near the cut-off rate $R_0 = 4.0$ bits/dimension, at present). Some bounds presented here have been applied to the Leech lattice at rate $R = 7$ bits/2dimensions. These results show that in spite of the very high number of neighbours, P_E may be quite acceptable. Both near and above the cut-off rate, the union bound is useless.

III.1.6. CONCLUSIONS

The Minkowski-Hlawka theorem of the geometry of numbers was used by R. de Buda in place of the random coding argument to show that lattice codes are close to the optimal codes at high signal-to-noise ratios. This result was modified here to prove

capacity theorem for the Gaussian channel (negative and positive statements) when high-dimensional lattices are used as signal constellations. This approach allows a better understanding of the sphere hardening and the use of lattice codes to achieve the channel capacity for any SNR. Error bounds are also presented, including an $E(R, C)$ reliability function for lattice coded AWGN channels.

REFERENCES

- [1] C.E. Shannon, Probability of error for optimal codes in a Gaussian channel, *Bell Syst. Tech. J.*, **38**, 1959, pp. 611-656
- [2] G.D. Forney Jr., Coding and its application in space communications, *IEEE Spectrum*, **7**, Jun., 1970, pp. 47-58
- [3] G. Ungerboeck, Channel coding with multilevel/phase signals, *IEEE Trans. Info. Theory*, **IT-28**, n.1, Jan., 1982, pp. 56-67
- [4] A.R. Calderbank, The mathematics of MODEMs, *Math. Intell.*, to appear
- [5] I.F. Blake, The Leech lattice as a code for the Gaussian channel, *Infor. Contr.*, **19**, 1971, pp. 66-74
- [6] J. Leech and N.J.A. Sloane, Sphere packings and error correcting codes, *Canad. J. Math.*, **23**, 1971, pp. 71-745
- [7] N.J.A. Sloane, Tables of sphere packings and spherical codes, *IEEE Trans. Info. Theory*, **IT-27**, n.3, May, 1981, pp. 327-338
- [8] A.R. Calderbank and N.J.A. Sloane, New trellis codes based on lattices and cosets, *IEEE Trans. Info. Theory*, **IT-33**, n.2, Mar., 1987, pp. 177-195
- [9] G.D. Forney Jr., Coset codes. part I: Introduction and geometrical classification, *IEEE Trans. Info. Theory*, **IT-34**, n.5, Sept., 1988, pp. 1123-1151
- [10] G.D. Forney Jr., Coset codes. part II: Binary lattices and related codes, *IEEE Trans. Info. Theory*, **IT-34**, n.5, Sept., 1988, pp. 1152-1187
- [11] R. de Buda, The upper error bound of a near-optimal code, *IEEE Trans. Info. Theory*, **IT-21**, n.4, July, 1975, pp. 441-445
- [12] R. de Buda, Some optimal codes have structure, *IEEE J. Select. Areas Comm.*, **SAC-7**, n.6, Aug., 1989, pp. 893-899
- [13] J.H. Conway and N.J.A. Sloane, Sphere packings, lattices and groups, NY:Springer, 1988
- [14] M. Abramowitz and I.A. Stegun Eds., Handbook of mathematical functions, Washington, DC:National Bureau of Standards, 1972
- [15] G.R. Lang and F.M. Longstaff, A Leech lattice MODEM, *IEEE J. Select. Areas Comm.*, **SAC-7**, n.6, Aug., 1989, pp. 968-973
- [16] Y. Be'ery, B. Shahar and J. Snyders, Fast decoding of the Leech lattice, *IEEE J. Select. Areas Comm.*, **SAC-7**, n.6, Aug., 1989, pp. 959-967

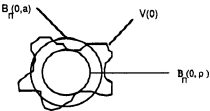


Figure 1. Regions concerning the bounds.

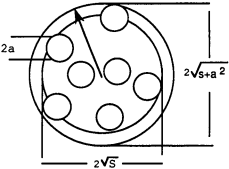


Figure 2. Sphere packing representation of lattice points.

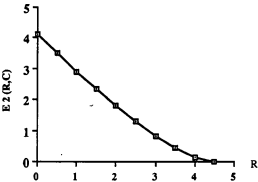


Figure 3. Reliability function $E_2(R, C)$ for a Gaussian channel ($\text{SNR} = 27.1 \text{ dB}$).

III.2. PERFORMANCE OF LATTICE CODES OVER THE GAUSSIAN CHANNEL

"Lattices are to block codes as classical music
is to rock and roll", N.J.A. SLOANE.

ABSTRACT

We develop an upper bound on the error probability of lattice codes combined with Quadrature Amplitude Modulation (QAM) over the additive white Gaussian channel. This bound depends on a lattice figure of merit and is readily put in exponential form by using Chernoff bound. An interesting lower bound is derived by a similar reasoning. We also discuss the continuous approximation used to estimate the average power normalized to two-dimensions, and suggest to improve it by using the sphere packing idea. Examples of performance evaluation are given for a few lattices. Finally, we present upper and lower bounds on the best fundamental coding gain per dimension (due both to density and covering) for an arbitrarily large number of dimensions. It is shown in the Appendix that, as the Ungerboeck codes, the lattice codes do not shape the signal power spectrum.

III.2.1. INTRODUCTION

There is a renewed interest in coded signaling strategies based on finite-dimensional constellations (e.g., lattices) in order to improve the noise immunity [1-6]. Ever since Shannon's monumental work it was clear that a strong relationship exists between increasing transmission rate toward capacity and achieving dense structures in high-dimensional spaces. However, most coding theorists have raised the interest in lattice codes just a while ago when solid bases of this theory were laid by Forney [7]. Furthermore, it has been recently reported that lattices exist which can achieve the capacity over the Gaussian channel [8,9] showing that this coding technique is quite powerful.

We deal in this paper with finite-dimensional signaling schemes based on lattices even though the ones supplied by (time-invariant) infinite-dimensional packings i.e., of the trellis-coded modulation type, appear to be still more powerful. We intend essentially to clarify the tradeoff between signal-to-noise ratio and transmission rate for such schemes. A constellation (an n -dimensional signal set), denoted by $\Omega = C(\Lambda, \mathbf{R})$, consists of the points of a lattice Λ (or a translate of Λ) that lie within a region \mathbf{R} . If $|\Omega|$ is the cardinality of the set Ω (i.e., the number of its signal points), then the transmission rate per dimension is given by

$$R = \frac{1}{n} \log_2 |\Omega| \quad \text{bits/dim.} \quad (1)$$

Furthermore, let $P(\Omega)$ denote the average power of the n -dimensional signal points (the average squared Euclidean norm of the constellation). The average power normalized to 2-dimensions (2D) is, therefore, given by

$$P(C) = \frac{2}{n} P(\Omega). \quad (2)$$

We shall develop upper and lower bounds on the (block) error probability which depend upon lattice parameters. Our interest will be focused on combining lattice coding and modulation. Examples of performance evaluation are provided for a few well-known lattices. We also discuss the asymptotic behaviour (for an arbitrarily large number of dimensions) of these bounds and we are led to conjecture that "covering" becomes more important than "packing" for rates beyond the cutoff rate. Our purpose is to prompt engineers to think of the deep connection between sphere packing problems and capacity, a very old matter. The last section provides some insightful arguments in this direction.

III.2.2. BOUNDS ON THE n -DIMENSIONAL ERROR PROBABILITY

We shall assume QAM modulation, which implies transmitting a 2D-symbol (phase and quadrature components) per time interval T . The corresponding waveform is

$$s_T(t) = a_c u(t) \cos(2\pi f_c t) - a_s u(t) \sin(2\pi f_c t), \quad (3)$$

where $u(t)$ is a narrow-band waveform, f_c is the carrier frequency, and (a_c, a_s) are the symbol coordinates. We assume that $n/2$ successive pairs of such symbol coordinates (n even) are restricted to be the coordinates of a point of an n -dimensional lattice. We then establish the following result:

THEOREM 1. The error probability of lattice codes over the Gaussian channel is upper bounded by $P_e(n) \leq Q(FM(\rho, \Omega) \gamma_{av} n | n)$, where $Q(\chi^2 | n)$ is the χ^2 -function with n degrees of freedom and $FM(\rho, \Omega)$ is the constellation figure of merit defined by (19), and γ_{av} is the signal-to-noise ratio per dimension.

proof. Supposing a signaling rate of $1/T$ bauds, the lattice coded QAM signal $s(t)$, in terms of its complex envelope, is given by

$$s(t) = \text{Re} \left\{ \sum_{k=-\infty}^{+\infty} \sum_{m=1}^{n/2} \underline{P}_k^{(m)} u \left[t - \left(k \frac{n}{2} + m \right) T \right] e^{j2\pi f_c t} \right\}, \quad (4)$$

where the complex $\underline{P}_k^{(m)} := a_{ck}^{(m)} + j a_{sk}^{(m)} \equiv (a_{ck}^{(m)}, a_{sk}^{(m)})$ represents a constituent 2D point or symbol.

The k -th transmitted signal corresponds to a lattice point $\vec{P}_k \in \Lambda$ whose coordinates are given by $n/2$ two-dimension constituents, that is

$$\vec{P}_k = (\underline{P}_k^{(1)}; \underline{P}_k^{(2)}; \dots; \underline{P}_k^{(n/2)}) \quad (5)$$

or, more explicitly, each coded signal is an n -dimensional point

$$\vec{P}_k = (a_{ck}^{(1)}, a_{sk}^{(1)}; a_{ck}^{(2)}, a_{sk}^{(2)}; \dots; a_{ck}^{(n/2)}, a_{sk}^{(n/2)}) . \quad (6)$$

It is shown in the Appendix that, as the Ungerboeck codes [10], the lattice codes do not shape the signal power spectrum. The coded signal has a power spectral density expressed by

$$S_s(f) = \frac{P(C)}{4T} \left\{ |U(f-f_c)|^2 + |U(-f-f_c)|^2 \right\} . \quad (7)$$

The average power, P_{av} , of the transmitted signal (*) is found by integrating the above equation, namely,

$$P_{av} = \int_{-\infty}^{\infty} S_s(f) df = \frac{P(C)}{T} \frac{\xi_u}{2} , \quad (8)$$

where $\xi_u := \int_{-\infty}^{+\infty} |u(t)|^2 dt = \int_{-\infty}^{+\infty} |U(f)|^2 df$ is the energy of the narrow-band pulse $u(t)$. Hence, the average energy of the transmitted signal $s(t)$ is $E_{av} = P(C) \xi_u / 2$.

On the other hand, the noise $n_w(t)$, is assumed to be an *additive white Gaussian* random process and we will use its pass-band representation in a bandwidth B hertz. Thus, the narrow-band waveform of the noise can be expressed as

$$n_w(t) = n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t), \quad (9)$$

where $n_c(t)$ and $n_s(t)$ are both low-pass noise signals band-limited to $B/2$ having as power spectral density

$$S_{nc}(f) = S_{ns}(f) = \begin{cases} N_0 & |f| < B/2; \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Moreover, it is assumed that the constituent 2D signals on the phase and quadrature carriers are perfectly separated by coherent detection and that maximum likelihood decoding (MLD) is carried out by the decoder after the reception of an n -dimensional signal (block). The diagram of such a receiver is shown in Figure 1. The received signal $s(t) + n_w(t)$ is processed as indicated: a coherent detection is performed, followed by matched filtering and decision-making after each block.

Following comments made in [11], we are interested in the signal-to-noise ratio per dimension:

(*) It should not be confounded with $P(C)$ which depends only on the lattice.

$$\gamma_{av} := E_{av}/N_0. \quad (11)$$

Each time interval T , a corrupted constituent 2D point $\underline{Y}_k^{(m)} = (y_{ck}^{(m)}, y_{sk}^{(m)})$ is received namely,

$$\underline{Y}_k^{(m)} = \underline{P}_k^{(m)} + \underline{W}_k^{(m)}. \quad (12)$$

Here, $\underline{W}_k^{(m)}$ is the constituent 2D noise vector which is given by $\underline{W}_k^{(m)} = (w_{ck}^{(m)}, w_{sk}^{(m)})$. Moreover, the random variables $w_{ck}^{(m)}$ and $w_{sk}^{(m)}$ are independent and identically distributed, namely Gaussian variables with zero mean and variance $\sigma^2 = N_0/\xi_{sm}$.

Thus, the received (noisy) n -dimensional vector, $\underline{\vec{Y}}_k = (\underline{Y}_k^{(1)}; \underline{Y}_k^{(2)}; \dots; \underline{Y}_k^{(n/2)})$, is expressed as

$$\underline{\vec{Y}}_k = \underline{\vec{P}}_k + \underline{\vec{W}}_k, \quad (13)$$

that is, the transmitted lattice point is perturbed by the n -dimensional noise vector:

$$\underline{\vec{W}}_k = (\underline{W}_k^{(1)}; \underline{W}_k^{(2)}; \dots; \underline{W}_k^{(n/2)}). \quad (14)$$

Furthermore, it is well known that MLD on the Gaussian channel is equivalent to choose the lattice point closest to the received signal, that is, decision regions are exactly the Voronoi regions. The minimum Euclidean distance between two lattice points is $d_{\min} = 2\rho$. A simple bound on the block error probability (an n -dimensional received point) can be developed taking

$$P_e(n) \leq P(|\underline{\vec{W}}_k| \geq d_{\min}/2). \quad (15)$$

The squared norm of the noise vector can easily be found using (14),

$$|\underline{\vec{W}}_k|^2 = \sum_{m=1}^{n/2} [W_{ck}^{(m)}]^2 + [W_{sk}^{(m)}]^2. \quad (16)$$

Therefore, we have the sum of n squared Gaussian random variables with the same distribution. This can be normalized in order to obtain an n -variate chi-square distribution [12] by putting (15) in the form

$$P_e(n) \leq P\left(\sum_{i=1}^n z_i^2 \geq \frac{d_{\min}^2/4}{\sigma^2}\right), \quad (17)$$

where the n new random variables z_i are normal ones with zero mean and variance unity. Of course, this is readily written in terms of the χ^2 -function with n degrees of freedom, resulting in

$$P_e(n) \leq Q\left(\frac{d_{\min}^2/4}{P(\Omega)} \gamma_{av} n | n\right), \quad (18)$$

where $Q(X|n) = \int_X^{\infty} \frac{x^{n/2-1} \exp(-x/2)}{2^{n/2} \Gamma(n/2)} dx$. Finally, a constellation figure of merit [9,11 also 13] has been defined to be

$$FM(\rho, \Omega) := \frac{\rho^2}{P(\Omega)} = \frac{d_{\min}^2/4}{P(\Omega)}, \quad (19)$$

and the proof follows. ■

If the *covering radius* r of the lattice is used in (15) instead of its *packing radius* ρ , the inequality is reversed and we get the following corollary:

COROLLARY. The error probability of lattice codes over the Gaussian channel is lower bounded by $P_e(n) \geq Q(FM(r, \Omega) \gamma_{av} n \mid n)$, where $FM(r, \Omega)$ is the constellation figure of merit :

$$FM(r, \Omega) := \frac{r^2}{P(\Omega)} \quad \blacksquare \quad (20)$$

For a very large number of dimensions, the upper bound is roughly given by [12]

$$P_e(n) \leq 1/2 \operatorname{erfc}(\sqrt{n}[\sqrt{FM(\rho, \Omega)\gamma_{av}} - 1]). \quad (21)$$

It is convenient to define a parameter λ^2 as

$$\lambda^2 := FM(\rho, \Omega) \gamma_{av}. \quad (22)$$

Obviously, for a fixed signal-to-noise ratio, this parameter λ^2 depends on the transmission rate R .

In the following, we present some applications of the above Theorem. First we remark that a signal-to-noise ratio $\gamma_{av} = 22$ decibels is a quite high SNR for a 16-point QAM constellation while it is a very poor one for a 256-ary constellation. A simple question arises: what exactly does a high SNR mean? Of course, a good definition should depend on the constellation (in particular on its dimensionality). We have used the bound of Theorem 1 for the purpose of estimating the range of SNR resulting in an error probability within a given level ε , that is, according to

$$Q(FM(\rho, \Omega) \gamma_{av} n \mid n) \leq \varepsilon, \quad (23)$$

and we have plotted, in Figure 2, the values of λ^2 needed to guarantee an error rate less than $\varepsilon = 10^{-5}$. Thus, uncoded systems corresponding to M -ary QAM, must work at least at $3\gamma_{av}/2(M-1) > 11.5$, from which it follows that $\gamma_{av} > 7.67(M-1)$ in order to ensure an error probability less than 10^{-5} .

Additionally, it should also be pointed out that for a very large number of dimensions, the behaviour of the error probability does not depend on the number of neighbours at the minimum distance.

THEOREM 2. The n -dimensional error probability of a lattice-coded signal set is exponentially upper bounded by $P_e(n) \leq \exp[-n C(\lambda^2)]$,

$$\text{where } C(\lambda^2) = \begin{cases} 1/2(\lambda^2 - 1 - \ln \lambda^2) & \text{if } \lambda^2 \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

proof. We recall that the error probability is bounded according to Theorem 1 which can be rewritten as

$$P_e(n) \leq P\left(\sum_{i=1}^n z_i^2 \geq n\lambda^2\right). \quad (24)$$

Of course, the random variables $\{z_i\}$ are independent and identically distributed (i.i.d.) so Chernoff's bound [14] can readily be applied, resulting in

$$P_e(n) \leq \exp\{-n[s\lambda^2 - \ln M_{z^2}(s)]\} \quad \forall s \geq 0, \quad (25)$$

where $M_{z^2}(s) = (1-2s)^{-1/2}$ is the generator function of a z_i^2 variable. The tightest upper bound according to (25) is thus:

$$P_e(n) \leq \exp[-n C(\lambda^2)] \quad (26)$$

where

$$C(\lambda^2) := \max_{s \geq 0} [s\lambda^2 - \ln M_{z^2}(s)] := \max_{s \geq 0} F_\lambda(s). \quad (27)$$

It can easily be verified that $F_\lambda(s)$ is a convex upward function (convex \cap) and that its maximum occurs at $s = 1/2(\lambda^2 - 1)/\lambda^2$. Thus, two cases are to be considered depending on the transmission rate. First, if λ^2 is less than or equal to unity, we find immediately that $C(\lambda^2) = 0$. Second, we remark that $F_\lambda(s)$ has a maximum for $s > 0$ when λ^2 is greater than unity. Additionally, it follows from the fundamental inequality of information theory [14] (i.e., which expresses the convexity of the logarithmic function) that $C(\lambda^2)$ is strictly positive when the maximum is reached, which completes the proof. ■

The main results of this section can be summarized by

$$Q(FM(r, \Omega) \gamma_{av} n \mid n) \leq P_e(n) \leq Q(FM(\rho, \Omega) \gamma_{av} n \mid n), \quad (28)$$

where $\rho < r$ which implies $FM(\rho, \Omega) < FM(r, \Omega)$.

III.2.3. LATTICE POINTS CONFINED IN BOUNDED REGIONS

Forney and Wei [13] showed that the normalized power, $P(C)$, is approximately equal to the average power $P(R)$ of a continuous distribution that is uniform within R and zero elsewhere, resulting in

$$P(C) \approx 2G(R) V(R)^{2/n}, \quad (29)$$

where

$$V(R) = \int_R dv \quad \text{and} \quad G(R) = \frac{\int_R ||r||^2 dv}{n V(R)^{1+2/n}}. \quad (30)$$

Here, $V(\mathbf{R})$ is the volume of the region \mathbf{R} and $G(\mathbf{R})$ is its dimensionless second moment [15] which measures the effect of the shape of the region \mathbf{R} on the average signal power.

Forney and Wei (see [13] proposition 1, p. 883) also proposed the approximation

$$V(\mathbf{R}) \approx |\Omega| V(\mathbf{A}) \quad (31)$$

so (29) and (1) imply that the normalized power $P(C)$ is related to both the transmission rate and the fundamental volume. However, such an approach does not take into account the possibility of including boundary points. In the following, we go further in order to allow the use of the border of \mathbf{R} .

In order to get insight, we first consider the straightforward case of a 2-dimensional constellation with *power limitation per one dimension* according to Figure 3. The available volume is then $V(\mathbf{R}) = (2d_0)^2 = 4d_0^2$ while the fundamental volume (volume per signal point) is $V(\mathbf{A}) = d_0^2$. Therefore the total number of points, $|\Omega|$, that can be put into such a region is given by $V(\mathbf{R}) \approx |\Omega| V(\mathbf{A})$ or $|\Omega| = 4$.

However, it can clearly be seen from Figure 3.b that it is possible to put nine points inside the same region provided that boundary points are allowed. The counting could be done in the following way: we consider an augmented region \mathbf{R}^* , $V(\mathbf{R}^*) > V(\mathbf{R})$, and we use the relation $V(\mathbf{R}^*) \approx |\Omega| V(\mathbf{A})$. The volume of the augmented region is chosen such that the normalized volume is increased by that of the fundamental volume, so that

$$V(\mathbf{R}^*) = [V(\mathbf{R})^{1/2} + V(\mathbf{A})^{1/2}]^2,$$

where $V(\mathbf{R})^{1/2}$ and $V(\mathbf{A})^{1/2}$ are respectively the volume of \mathbf{R} per dimension and the additional volume per dimension. Therefore $V(\mathbf{R}^*) = (2d_0 + d_0)^2 = 9d_0^2$ so one has $|\Omega| = 9$.

One can indeed justify the approximation (31) in the following way. Consider a finite and full set Ω of points belonging to a lattice Λ , where "full" is intended here to mean that Ω contains contiguous points of the lattice and moreover that no holes are left i.e., any lattice point which is surrounded by points belonging to Ω also belongs to it.

We may *define* the region \mathbf{R} by associating with each point of Ω a fundamental region which contains it (i.e., whose volume is the fundamental volume $\det(\Lambda)$; its shape is definite but arbitrary). Then, the union of the fundamental regions clearly defines a region \mathbf{R} such that (31) obtains with equality. The fundamental regions associated with the lattice points tessellate the region \mathbf{R} .

This way of defining \mathbf{R} obviously results in the points of Ω being *inside* \mathbf{R} since each of them is surrounded with its own fundamental region. One may ask what is the number of points which can be put inside \mathbf{R} and on its border. One is thus naturally led to consider an augmented region \mathbf{R}^* which is the union of the fundamental regions associated with all the points inside \mathbf{R} and on its border. The inclusion of such an

additional volume results in

$$V(\mathbf{R}^*) \approx [V(\mathbf{R})^{2/n} + V(\mathbf{A})^{2/n}]^{n/2}.$$

The newly defined region \mathbf{R}^* now contains as many lattice points as \mathbf{R} has inside it and on its border and (31) will be applied to \mathbf{R}^* instead of \mathbf{R} . Stated a bit more formally:

PROPOSITION 1. (COUNTING): The size $|\Omega|$ of any constellation consisting of the points of a lattice Λ (or a translate of Λ) that lie within any bounded region \mathbf{R} of volume $V(\mathbf{R})$ is approximately

$$|\Omega| \approx \left\{ 1 + \left[\frac{V(\mathbf{R})}{V(\mathbf{A})} \right]^{2/n} \right\}^{n/2}. \quad (32)$$

proof. Consider two spherical regions $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{A}}$ related to \mathbf{R} and \mathbf{A} , respectively, in the following way: $\tilde{\mathbf{R}}$ (resp. $\tilde{\mathbf{A}}$) is defined as an n -sphere of volume equal to $V(\mathbf{R})$ (resp. $V(\mathbf{A})$). In order to include signal points belonging to the border, we introduce a new region $\tilde{\mathbf{R}}^*$, namely the augmented region, which contains the given region \mathbf{R} and such that the added normalized power is that of the fundamental volume. Since the normalized power is additive, we have

$$P(\tilde{\mathbf{R}}^*) = P(\tilde{\mathbf{R}}) + P(\tilde{\mathbf{A}}). \quad (33)$$

Applying the continuous approximation and since the dimensionless second moment is the same for the above regions, it follows that

$$V(\tilde{\mathbf{R}}^*)^{2/n} = V(\tilde{\mathbf{R}})^{2/n} + V(\tilde{\mathbf{A}})^{2/n}. \quad (34)$$

We now estimate the number of signal points in the same way as Forney did but using the augmented region $\tilde{\mathbf{R}}^*$ instead of \mathbf{R} , resulting in

$$V(\tilde{\mathbf{R}}^*) \approx [V(\mathbf{R})^{2/n} + V(\mathbf{A})^{2/n}]^{n/2} \approx |\Omega| V(\mathbf{A}), \quad (35)$$

and the proof follows. ■

The volume of a hypersphere of radius r is given by $J_n(r) = V_n r^n$, where $V_n = \pi^{n/2}/\Gamma(n/2+1)$ is the volume of an n -sphere of radius unity. Furthermore, suppose that the average signal power is S and the average noise power is N , considering both the signal and the noise as n -dimensional signals. If we take the radius of the region \mathbf{R} as $\tilde{r} = \sqrt{S}$ and the radius of the region \mathbf{A} as \tilde{a} , then the radius \tilde{r}^* of the augmented region \mathbf{R}^* can be found using (34), that is,

$$J_n(\tilde{r}^*)^{2/n} = J_n(\sqrt{S})^{2/n} + J_n(\tilde{a})^{2/n}, \quad (36)$$

so that $\tilde{r}^* = \sqrt{S + \tilde{a}^2}$. Therefore, (35) can be rewritten as

$$J_n(\tilde{r}^*) \approx |\Omega| J_n(\tilde{a}). \quad (37)$$

This results in a transmission rate of $R = 1/2 \log_2(1 + S/\tilde{a}^2)$. This kind idea has been

used in conjunction with the sphere hardening argument (one must have $\tilde{a}^2 > N$ as $n \rightarrow \infty$ in order to assert that the error probability goes to zero) in the original Shannon's proof of capacity theorem [16] and recently in the proof that lattices exist which can achieve the channel capacity [9].

PROPOSITION 2. (CONTINUOUS APPROXIMATION): The average signal power normalized to 2 dimensions of any constellation Ω consisting of the points of a lattice Λ (or a translate of Λ) that lie within a region \mathbf{R} is approximately equal to

$$P(\mathbf{C}) \approx 2G(\mathbf{R}) [2^{2R} - 1] V(\Lambda)^{2/n}, \quad (38)$$

where R is the transmission rate in bits/dim,

$G(\mathbf{R})$ is the normalized or dimensionless second moment,

$V(\Lambda)$ is the fundamental volume of Λ .

proof. The proof follows by applying proposition 1 combined with (1) in order to estimate $V(\mathbf{R})$ and then substituting the result in (29). ■

Next, we apply the above proposition in order to estimate the normalized average signal power of simple lattice-based constellations. First, we consider the M -point monodimensional amplitude signaling (PAM). The signal points are M odd-integers namely $\pm(2i+1)$, $i = 0, 1, \dots, M/2-1$, resulting in the true average power [17]

$$P(\mathbf{C}) = 2(M^2-1)/3. \quad (39)$$

On the other hand, this signaling scheme corresponds to take all the points belonging to the translate $2\mathbf{Z}+1$ of the lattice $\Lambda = 2\mathbf{Z}$ (that is, the one dimensional even-integer grid) and a region $\mathbf{R} = [-(M-1), (M-1)]$. Thus, one has $V(\Lambda) = 2$ and $V(\mathbf{R}) = 2(M-1)$, and the corresponding normalized second moment is

$$G(\mathbf{R}) = \frac{1}{1 [2(M-1)]^3} \cdot \int_{-(M-1)}^{(M-1)} v^2 dv = \frac{1}{12}, \quad (40)$$

so proposition 2 yields $P(\mathbf{C}) \approx 2.1/12.[M^2-1].4$ i.e., the exact result (39).

In a second closely related example, we consider the M -point two-dimensional narrow sense QAM. Now, the signal set is constituted by all the points in the odd-integer grid (the translate $2\mathbf{Z}^2 + (1,1)$ of the lattice $2\mathbf{Z}^2$) within $\mathbf{R} = [-\sqrt{M-1}, \sqrt{M-1}]^2$. Remembering that the normalized second moment of an n -cube is $1/12$, and that $V(\Lambda) = 4$, we get

$$P(\mathbf{C}) \approx 2.1/12.[M-1].4 = 2(M-1)/3. \quad (41)$$

III.2.4. GAINS, VORONOI REGIONS AND MORE ON THE BEHAVIOUR OF LATTICES

The gain $\gamma(\Omega_1; \Omega_2)$ of a constellation $\Omega_1 = C(\mathbf{R}_1, \Lambda_1)$ with respect to another constellation $\Omega_2 = C(\mathbf{R}_2, \Lambda_2)$ at the same rate R can be measured by the ratio of their figures of merit, the superscript minus (plus) standing for values based on the *inradius* ρ (*circumradius* r) of a Voronoi region.

$$\gamma^-(\Omega_1; \Omega_2) := \frac{FM(\rho, \Omega_1)}{FM(\rho, \Omega_2)} = \frac{\rho_1^2}{\rho_2^2} \cdot \frac{P(C_2)}{P(C_1)}, \quad (42.a)$$

$$\gamma^+(\Omega_1; \Omega_2) := \frac{FM(r, \Omega_1)}{FM(r, \Omega_2)} = \frac{r_1^2}{r_2^2} \cdot \frac{P(C_2)}{P(C_1)}. \quad (42.b)$$

Now, under the continuous approximation (38), it follows that

$$\gamma^-(\Omega_1; \Omega_2) = \frac{\rho_1^2}{\rho_2^2} \cdot \frac{G(\mathbf{R}_2)}{G(\mathbf{R}_1)} \cdot \frac{V(\Lambda_2)^{2/n}}{V(\Lambda_1)^{2/n}}, \quad (43.a)$$

$$\gamma^+(\Omega_1; \Omega_2) = \frac{r_1^2}{r_2^2} \cdot \frac{G(\mathbf{R}_2)}{G(\mathbf{R}_1)} \cdot \frac{V(\Lambda_2)^{2/n}}{V(\Lambda_1)^{2/n}}. \quad (43.b)$$

We can, therefore, split the total gains in two parts: the coding gains γ_c and the shape gains γ_s , given respectively by

$$\gamma_c^-(\Lambda_1; \Lambda_2) := \frac{\delta^{2/n}(\Lambda_1)}{\delta^{2/n}(\Lambda_2)}, \quad (44.a)$$

$$\gamma_c^+(\Lambda_1; \Lambda_2) := \frac{\theta^{2/n}(\Lambda_1)}{\theta^{2/n}(\Lambda_2)}, \quad (44.b)$$

where δ and θ are the center density and the normalized thickness of the lattice i.e., $\delta := \frac{\rho^n}{V(\Lambda)}$ and $\theta := \frac{r^n}{V(\Lambda)}$ (see e.g., [18, p. 13 and p. 32, resp.]), and

$$\gamma_s^-(\mathbf{R}_1; \mathbf{R}_2) = \gamma_s^+(\mathbf{R}_1; \mathbf{R}_2) := \frac{1/G(\mathbf{R}_1)}{1/G(\mathbf{R}_2)}. \quad (45)$$

The coding gain is easily seen to be the power saving due to the thickness of the lattice employed. Thus, the gain as defined in (43) can be expressed as

$$\gamma^-(\Omega_1; \Omega_2) = \gamma_c^-(\Lambda_1; \Lambda_2) \gamma_s^-(\mathbf{R}_1; \mathbf{R}_2), \quad (46.a)$$

$$\gamma^+(\Omega_1; \Omega_2) = \gamma_c^+(\Lambda_1; \Lambda_2) \gamma_s^+(\mathbf{R}_1; \mathbf{R}_2). \quad (46.b)$$

In the particular case where the reference constellation is the uncoded one (i.e., that based on the cubic lattice Z^n), we derive the expressions

$$\gamma_c^-(\Lambda: Z^n) := \gamma_c^-(\Lambda) = 48^{2/n}(\Lambda), \quad (47.a)$$

$$\gamma_c^+(\Lambda: Z^n) := \gamma_c^+(\Lambda) = 2\theta^{2/n}(\Lambda), \quad (47.b)$$

and

$$\gamma_s(\mathbf{R}_1: n\text{-cube}) := \gamma_s(\mathbf{R}) = \frac{1}{12G(\mathbf{R})}. \quad (48)$$

We remark yet that the coding gain is

$$\gamma_c^-(\Lambda) = 4\rho^2/\det^{2/n}(\Lambda). \quad (49)$$

We can rewrite the above equation keeping Forney's notation, recalling that the fundamental volume is $V(\Lambda) = \det \Lambda$ and that $\rho = d_{\min}/2$. Thus, the coding gain is $\gamma_c^-(\Lambda) = d_{\min}^2/V(\Lambda)^{2/n}$ which agrees with the definition of the fundamental gain [7].

The application of the above approximations to Theorem 1 gives the following upper bound on the error probability:

$$P_e(n) \leq Q(\gamma^-(\Lambda: Z^n) FM(\rho, Z^n) \gamma_{sv} n | n) \quad (50)$$

i.e., explicitly:

$$P_e(n) \leq Q\left(\frac{\gamma_c^-(\Lambda)\gamma_s^-(\mathbf{R})}{n/2} \frac{3\gamma_{sv}}{2[2^{2R}-1]} n | n\right). \quad (51)$$

The above bound has been applied to a few lattices for different transmission rates in bits per two-dimension. Figure 4 presents performance curves which have been evaluated for a few known lattices over the Gaussian channel (as well as the performance of the uncoded related systems). In order to compare the performance of signal sets, we must use a common criterion for the error rate. Since our bounds concern the n -dimensional error probability, we should evaluate the performance of the uncoded system taking into account n dimensions. In this event, we have $n/2$ independent 2-dimensional constituent signals (n even), so that the error rate is

$$P_e(n)_{\text{unc}} = 1 - [1 - P_e(2)_{\text{unc}}]^{n/2}. \quad (52)$$

Furthermore, the symbol (2-dimensional) error rate for the uncoded system can be upper bounded by Theorem 1 so as to allow a fair comparison, that is

$$P_e(n)_{\text{unc}} \leq 1 - [1 - Q\left(\frac{3\gamma_{sv}}{2(2^{2R}-1)} 2|2| \right)^{n/2}]. \quad (53)$$

Similarly to (51), the following lower bound is easily derived:

$$P_e(n) \geq Q\left(\frac{\gamma_c^+(\Lambda)\gamma_s^+(\mathbf{R})}{n/2} \frac{3\gamma_{av}}{[2^{2R}-1]} n |n\right). \quad (55)$$

We also present bounds on the error rate for the Leech lattice compared with experimental results obtained from a commercial modem [19]. As it can be seen from Figure 6, the upper bound gives a fairly good agreement with the measured performance of the prototype (within -1.5 dB). Still, the lower bound calculated using the covering radius of the Leech packing [18, ch. 23], $r = \sqrt{2} \rho$, is remarkably tight. We believe that the tightness of the lower bound (based on the covering radius) may be explained by sphere hardening properties.

Incidentally, it should be stressed that modems based on the Leech lattice [e.g., 19] seem to be less attractive than those based on multidimensional Trellis-Coded Modulation [e.g. 25-27] owing to the high normalized redundancy (per two dimensional constellation) of such a lattice [7, table I]. Therefore, although excellent performance is obtained in the presence of white Gaussian noise, on real voiceband telephone channels things could not be so well because other impairments, like phase jitter, nonlinear distortion and frequency offset, become more and more important as the size of the two-dimensional constituent constellation increases. We now turn to the asymptotic behaviour of the lattices as n increases indefinitely.

PROPOSITION 3. (ASYMPTOTIC GAINS): The fundamental gain of the best lattice (covering) Λ for an arbitrarily large number of dimensions verifies the following bounds:

$$\frac{1}{2\pi e} \leq \lim_{n \rightarrow \infty} \frac{\gamma_c^-(\Lambda)}{n} < \frac{1}{\pi e}, \quad (56.a)$$

and

$$\lim_{n \rightarrow \infty} \frac{\gamma_c^+(\Lambda)}{n} = \frac{1}{\pi e}. \quad (56.b)$$

Moreover, the shape gain for any dimensionality is upper bounded by (Forney *et al.* [13,21])

$$1 \leq \lim_{n \rightarrow \infty} \gamma_s(\mathbf{R}) \leq \frac{\pi e}{6} \quad \blacksquare \quad (57)$$

proof. Bounds (56.a) are an immediate consequence of known bounds on the highest attainable density of lattices for arbitrary large dimensions. They follow from the Minkowski lower bound and the Blichfeldt (or Roeger) upper bound [18,20] and can be attributed to Conway and Sloane [18, p. 20]. These bounds read

$\exp_2(-n) \leq \Delta < \exp_2(-0.5n)$ for large n , $\exp_2(x) := 2^x$, where $\Delta = V_n \delta = V_n \rho^n / V(\Lambda)$. Since $\gamma_c^-(\Lambda) = 4\delta^{2/n}$, we obtain bounds on $\gamma_c^-(\Lambda)$ by raising the above bounds to the power $2/n$ and multiplying them by 4. This results in

$$1/V_n^{2/n} \leq \gamma_c^-(\Lambda) < 2/V_n^{2/n}.$$

But $V_n = \pi^{n/2} / \Gamma(n/2 + 1)$ so that Stirling's approximation results in

$$\frac{1}{n V_n^{2/n}} = \frac{1}{2\pi e} \frac{e^{-\Gamma^{2/n}(\frac{n}{2} + 1)}}{n/2} \rightarrow \frac{1}{2\pi e} \quad (58)$$

and (56.a) follows.

Similarly, $\gamma_c^+(\Lambda) = 2\theta^{2/n}$. The thickness of the lattice $\Theta = V_n \theta = \frac{V_n r^n}{V(\Lambda)}$ obeys the following bounds on the thinnest covering [18, p. 40] for large dimensionality:

$$n e^{-3/2} \leq \Theta \leq n \ln n + n \ln \ln n + 5n, \quad (59)$$

from which we derive

$$\frac{n^{2/n} e^{-3/2}}{V_n^{2/n}} \leq \theta^{2/n} \leq \frac{n^{2/n}}{V_n^{2/n}} [\ln n + \ln \ln n + 5]^{2/n}, \quad (60)$$

resulting in

$$\frac{2n^{2/n} e^{-3/2}}{n^{1/2/n}} \leq \frac{\gamma_c^+(\Lambda)}{n} \leq \frac{2n^{2/n}}{n^{1/2/n}} [\ln n + \ln \ln n + 5]^{2/n}. \quad (61)$$

The asymptotic behaviour for large n is expressed by (56.b), as a result of Stirling's approximation (58), which completes the proof. ■

III.2.5. BY WAY OF INTERPRETATION

We remark that although the code gain is limited to about 9 dB at the error probability level of 10^{-5} - 10^{-6} [16,21], we see clearly that the fundamental (asymptotic) gain is unbounded. Moreover, the equality $r^2/\rho^2 = \theta^{2/n}/\delta^{2/n}$ results in $r^2/\rho^2 = 2\gamma_c^+/\gamma_c^- > 2$, n large. We thus have two different topological aspects. We can "feel" a Voronoi region either touching it by the inside (increasing *inradius*, its maximum value being ρ) or touching it by the outside (decreasing *circumradius*, its minimum value being r). There is a fundamental discontinuity here: the upper and lower bounds (51) and (55) do not converge as dimensionality increases without limit.

The asymptotic behaviour of lattices can be examined using the following result [9]: $\lim_{n \rightarrow \infty} Q(\lambda^2 n | n) = 1$ or 0 according to whether λ^2 is less than unity or not, respectively. Supposing that the upper bound (56.a) could be achieved, we find that the

error probability for lattices with a large number of dimensions goes to zero exponentially for any rate $R < R_0 = 1/2 \log_2(1 + \gamma_{av}/2)$. However, this upper bound on the best (normalized) fundamental gain was improved by using Kabatiansky-Levenshtein-Sidel'nikov results [18, p.20] yielding $0.87176/\pi e$ instead of $1/\pi e$, which shows that even the cutoff rate cannot be attained according to this bound (it is too loose and is probably related to the critical rate ?). Constraints should therefore be relaxed in order to increase the transmission rate. This gives an additional evidence for the fact that the minimum distance criterion becomes irrelevant to very long (i.e., very-high dimensional) constellations [22,23] supporting high rates.

In contrast, good coverings at large dimensions are such that $\lim P_e(n) = 1$ provided that $R > C$, giving some kind of strong converse to the channel coding theorem. By allowing *partial* overlap when packing hyperspheres, we gain more degrees of freedom to construct signal sets and thus more centers per volume unit (higher rates) can be achieved. For instance, the decoding rule used in [9] to prove the existence of good lattices used as decision regions hyperspheres around each lattice point with partial overlap.

Even with some overlap (a bit less than covering), the error probability can not however be bounded away from zero under the sole assumption that $R < C$ [9,24]. We guess that the problem of achieving capacity is rather linked to the *asymptotic normalized thickness* for arbitrarily large dimensions than to the asymptotic center density as believed by many researchers.

Now we present in a relaxed and discursive style some heuristic geometrical considerations on very-high dimensional Euclidean spaces in relation with sphere hardening phenomena. Let us initially consider a hypersphere of radius a having the same volume as a Voronoi region, that is, $J_n(a) = V_n a^n = \det \Lambda$. Clearly, $\rho < a < r$ (ρ being the packing radius and r the covering radius). Let the decision regions be hyperspheres of radius s , $\rho < s < a < r$, surrounding each lattice point. Of course, these decision regions will overlap since $s > \rho$, but only partially since $s < r$. Therefore

$$\rho^2 < s^2 < \frac{\det^{2/n} \Lambda}{V_n^{2/n}} < r^2. \quad (62)$$

The asymptotic behaviour for large n can be examined using (58),

$$\frac{\rho^2}{\det^{2/n} \Lambda} < \frac{s^2}{\det^{2/n} \Lambda} \leq \frac{n}{2\pi e} \leq \frac{r^2}{\det^{2/n} \Lambda}, \quad (63)$$

which gives the bounds $\delta^{2/n}/n < 1/2\pi e \leq \theta^{2/n}/n$, n large. The tightness of the right-side bound is essentially due to the sphere hardening phenomenon. Almost the whole volume of the Voronoi region becomes concentrated near the surface, so that $a \leq r$. High-dimensional Voronoi regions behave *essentially* as spherical ones with radius

almost equal to the covering radius. In particular, given an $\epsilon > 0$, there exists n large enough such that $J_n(a) = J_n(r - \epsilon/2) < J_n(r)$. We know that covering exists such that this bound is tight (e.g., those for the best covering). Assume finally that $s = a - \epsilon/2$, so that $s = r - \epsilon$ (working a bit less than a covering), then **almost surely** the received point will be found inside the Voronoi region of the transmitted point provided that the (hardened) noise sphere is contained in the decision sphere. These ideas agree entirely with our previous results [22] and are illustrated in Figure 7. Interpreting capacity strictly as a "pure" sphere packing problem we should attain $\Delta^{2/n} \rightarrow 1$. Now, this is far from being true: no packing of hyperspheres without overlap does exist which covers the Euclidean space for any dimension $n > 1$. According to the above results, we propose to rather interpret a family of good codes (i.e., those which asymptotically achieve capacity) as a family of coverings for which $\Theta^{2/n} \rightarrow 1$ ($\theta^{2/n}/n \rightarrow 1/2\pi e$) despite the fact that $\Theta \rightarrow \infty$.

III.2.6. CONCLUSIONS AND SUMMARY

We have presented upper and lower bounds on the performance of communication systems based on lattice codes combined with QAM which are a function of constellation figures of merit. Our aim was to gain insight in the tradeoff between transmission rate and signal-to-noise ratio. We have also shown that lattice coding does not shape the power spectral density of the signal with respect to the uncoded QAM. Further, an improvement has been proposed on the continuous approximation by Forney. Bounds on the best fundamental gain per dimension (due to density and covering) have been presented for arbitrarily large dimensions. The chief value we would attack is the conjecture that the problem of achieving capacity is connected to the asymptotic behaviour of the normalized thickness of lattices.

APPENDIX: Power spectral density of lattice coded QAM modulation

Sketch of the proof. Denote by $s_b(t)$ the complex envelope of the QAM signal $s(t)$

$$s(t) = \text{Re} \{s_b(t) \exp(j2\pi f_c t)\}, \quad (\text{A1})$$

$$\text{where } s_b(t) := \sum_{k=-\infty}^{+\infty} \sum_{m=1}^{n/2} \underline{P}_k^{(m)} u_{k,m}(t), \quad u_{k,m}(t) := u(t - mT - k\frac{n}{2}T).$$

It can be shown that $s_b(t)$ is a cyclostationary process [28] but we treat it as a wide-sense stationary one by introducing a random phase uniformly distributed over one block. Therefore we consider a related signal

$$\tilde{s}_b(t) := \sum_{k=-\infty}^{+\infty} \sum_{m=1}^{n/2} \underline{P}_k^{(m)} u_{k,m}(t, \theta), \quad (\text{A2})$$

where $u_{k,m}(t, \theta) := u_{k,m}(t - \theta)$, θ being uniformly distributed in the interval $(0, \frac{n}{2}T)$.

The autocorrelation function (ACF) $R_{\tilde{s}_b}(t, t-\tau) = \mathbb{E}\{\tilde{s}_b(t) \tilde{s}_b^*(t-\tau)\}$ of the complex envelope is

$$R_{\tilde{s}_b}(t, t-\tau) = \sum_{k=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{m=1}^{\infty} \sum_{l=0}^{n/2} \mathbb{E} \underline{P}_k^{(m)} \underline{P}_{k-j}^{(m-l)*} \frac{2}{nT} \int_0^{\frac{n}{2}T} u_{k,m}(t, \theta) u_{m-l, k-j}^*(t-\tau, \theta) d\theta, \quad (\text{A3})$$

where $m-l$ is taken mod $n/2+1$.

We remark first that blocks are transmitted independently so that complex symbols on different blocks are uncorrelated, yielding $\mathbb{E} \underline{P}_k^{(m)} \underline{P}_{k-j}^{(m-l)*} = 0$, $j \neq 0$. Furthermore, we suppose the sequence of n -dimensional signals is wide-sense stationary and that there exists a block ACF

$$R_{\underline{P}}^{(l)} = \sum_{m=1, m \neq l}^{n/2} \mathbb{E} \{ \underline{P}_k^{(m)} \underline{P}_k^{(m-l)*} \}, \quad l = 0, 1, \dots, n/2. \quad (\text{A4})$$

Therefore, (A3) can be rewritten as

$$R_{\tilde{s}_b}(t, t-\tau) = \frac{2}{nT} \sum_{l=0}^{n/2} \sum_{m=1, m \neq l}^{n/2} \mathbb{E} \underline{P}_k^{(m)} \underline{P}_k^{(m-l)*} \sum_{k=-\infty}^{\infty} \int_0^{\frac{n}{2}T} u_{k,m}(t, \theta) u_{m-l, k}^*(t-\tau, \theta) d\theta. \quad (\text{A5})$$

By an appropriate change of variables, we obtain

$$R_{\tilde{s}_b}(t, t-\tau) = \frac{2}{nT} \sum_{l=0}^{n/2} \sum_{m=1, m \neq l}^{n/2} \mathbb{E} \underline{P}_k^{(m)} \underline{P}_k^{(m-l)*} \int_{-\infty}^{\infty} u(\alpha + \tau - lT) u^*(\alpha) d\alpha. \quad (\text{A6})$$

Putting the above equation in a simpler notation results in:

$$R_{\tilde{s}_b}(\tau) := R_{\tilde{s}_b}(t, t-\tau) = \frac{2}{nT} \sum_{l=0}^{n/2} R_{\underline{P}}^{(l)} v(\tau - lT), \quad (\text{A7})$$

where $v(\tau - lT) := \int_{-\infty}^{\infty} u(\alpha + \tau - lT) u^*(\alpha) d\alpha$.

Indeed, $R_{\underline{P}}^{(0)} = \sum_{m=1}^{n/2} \mathbb{E} |\underline{P}_k^{(m)}|^2 = P(\Omega)$. Moreover, it follows from lattice symmetry that $R_{\underline{P}}^{(l)} = 0$, $l \neq 0$. Taking now the Fourier transform of the ACF (A7), we have finally that

$$S_{\tilde{s}_b}(f) = \frac{P(\Omega)}{nT/2} |U(f)|^2 = \frac{P(C)}{T} |U(f)|^2, \quad (\text{A8})$$

so the power spectrum of the QAM lattice coded signal follows directly from the modulation theorem,

$$S_s(f) = \frac{1}{4} \{S_{\tilde{s}_b}(f-f_0) + S_{\tilde{s}_b}(-f-f_0)\}. \quad (A9)$$



REFERENCES

- [1] J. Leech, "Notes on Sphere Packing", *Canad. J. Math.*, **19**, 1967, pp. 251-267
- [2] J. Leech and N.J.A. Sloane, "Sphere Packing and Error-Correcting Codes", *Canad. J. Math.*, **XXII**, n 4, 1971, pp. 718-745
- [3] I.F. Blake, "The Leech Lattice as a Code for the Gaussian Channel", *Info. Contr.*, **19**, 1971, 66-74
- [4] N.J.A. Sloane, "A Note on the Leech Lattice as a Code for the Gaussian Channel", *Info. Contr.*, **46**, pp. 270-272, 1980
- [5] N.J.A. Sloane, "Tables of Sphere Packings and Spherical Codes", *IEEE Trans. Info. Theory*, **IT 27**, May 1981, pp. 327-338
- [6] A.R. Calderbank and N.J.A. Sloane, "New Trellis Codes Based on Lattice and Cosets", *IEEE Trans. Info. Theory*, **IT 33**, Mar. 1987, pp. 177-195
- [7] G.D. Forney Jr., "Coset Codes - part I: Introduction and Geometrical Classification", *IEEE Trans. Info. Theory*, **IT-34**, Sept. 1988, pp. 1123-1151 (1990 Donald G. Fink Prize)
- [8] R. De Buda, "Some Optimal Codes Have Structure", *IEEE J. Select. Areas on Commun.*, **SAC 7**, n.6, Aug. 1989, pp. 893-899
- [9] H. Magalhães de Oliveira and G. Battail, "A Capacity Theorem for Lattice Codes on Gaussian Channels", *proc. IEEE/SBT Int. Symp. Telecomm.*, Sept. 1990, Rio de Janeiro, Brazil, pp. 1.2.1-1.2.5
- [10] E. Biglieri, "Ungerboeck Codes Do not Shape the Signal Power Spectrum", *IEEE Trans. Info. Theory*, **IT 32**, July 1986, pp. 595-596
- [11] H. Magalhães de Oliveira and G. Battail, "On Generalized Cross Constellations and the Opportunistic Secondary Channel", submitted to *Ann. Telecomm.*
- [12] M. Abramowitz and I. Stegun, Eds., *Handbook of Mathematical Functions*, Washington, DC:NBS, 1970
- [13] G.D. Forney Jr. and L-F. Wei, "Multidimensional Constellations - part I: Introduction, Figures of Merit, and Generalized Cross Constellations", *J. Select. Areas Commun.*, **SAC 7**, Aug. 1989, pp. 877-892
- [14] R.E. Blahut, *Principles and Practice of Information Theory*, Reading, Massachusetts:Addison-Wesley, 1987

- [15] J.H. Conway and N.J.A. Sloane, "Voronoi Regions of Lattices, Second Moments of Polytopes, and Quantization", *IEEE Trans. Info. Theory*, IT-28, Mar. 1982, pp. 211-226
- [16] C.E. Shannon, "Communication in the Presence of Noise", in: Key papers in the Development of Info. Theory, D. Slepian (Ed.), reprinted from *proc. IRE*, 37, Jan., 1949, pp. 10-21
- [17] J.G. Proakis, *Digital Communications*, New York: McGraw-Hill, 2nd Ed., 1989
- [18] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, New York: Springer-Verlag, 1988
- [19] G.R. Lang and F.M. Longstaff, "A Leech Lattice modem", *IEEE J. Select. Areas Commun.*, SAC 7, Aug. 1988, pp. 968-973
- [20] N.J.A. Sloane, "The Packing of Spheres", *Scientific American*, 28, Jan. 1984, pp. 92-101
- [21] G.D. Forney Jr., R.G. Gallager, G.R. Lang, F.M. Longstaff and S.U. Qureshi, "Efficient modulation for band-limited channels", *IEEE J. Select. Areas Commun.*, SAC 2, 1984, pp. 632-647
- [22] H. Magalhães de Oliveira and G. Battail, "Random Coded Modulation: Performance and Euclidean distance spectrum evaluation", 1991 *IEEE Int. Symp. Info. Theory*, ISIT'91, Budapest, Hungary, June 1991.
- [23] G. Battail, "Construction explicite de bons codes longs", *Ann. Télécomm.*, 44, 1989, pp. 392-404
- [24] R. De Buda, "The upper Error Bound of a New Near-Optimal Code", *IEEE Trans. Info. Theory*, IT-21, July 1975, pp. 441-445
- [25] L-F. Wei, "Trellis-Coded Modulation with Multidimensional Constellations", *IEEE Trans. Info. Theory*, IT-33, n° 4, July 1987, pp. 483-501 (1989 Info. Theory Soc. IEEE Prize)
- [26] Y. Tanaka *et al.*, "A 19.2 kb/s High-Speed Voiceband Data Modem based on Eight-Dimensional Trellis Coding", *Globecom'87*, Tokyo, Japan, Nov. 1987, 2, pp. 36.1.1-36.1.5
- [27] G. Baudoin and M.S. Mitrani, "Modem à 19.200 bits/s sur Circuit Téléphonique", *Comm. & Trans.*, n° 2, 1989, pp. 73-84
- [28] W.A. Gardner and L.E. Franks, "Characterization of cyclostationary random processes", *IEEE Trans. Info. Theory*, IT-21, Jan. 1975, pp. 4-15

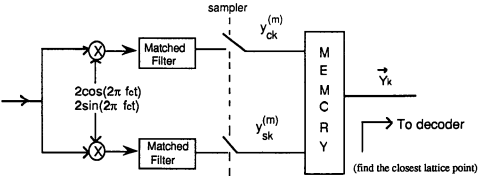


Figure 1. Coherent QAM receiver for lattice-coded signals

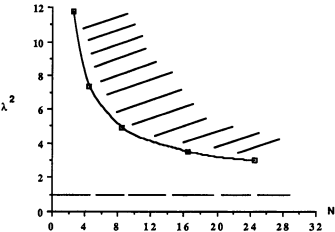


Figure 2. Lattice figure of merit times the signal-to-noise ratio required for keeping the error rate below 10^{-5} as a function of the dimensionality.

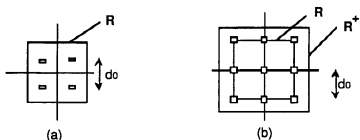
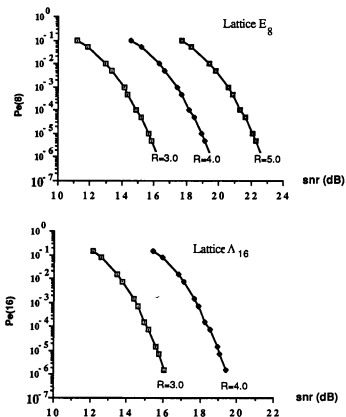


Figure 3. Putting signal points into a 2-dimensional region with power limitation per dimension. Fig. 3.a when only inner points are allowed, fig. 3.b if points can be chosen inside and on the border of the region.



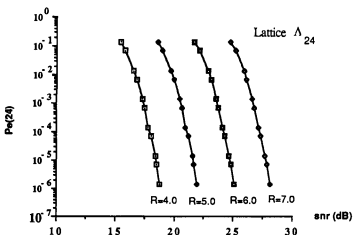


Figure 4. Bounds on the performance of lattices at several transmission rates in the presence of white Gaussian noise.

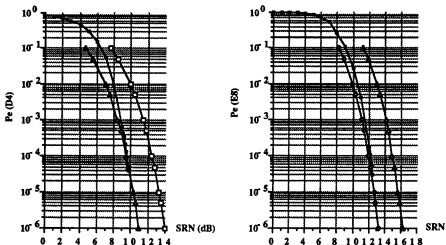


Figure 5. Upper and lower bounds on the block error probability for lattices over an AWGN channel: a) Schläfli lattice D_4 supporting 1.5 (bits/s)/Hz and b) Gosset lattice E_8 supporting 3.0 (bits/s)/Hz.

Simulation results have been obtained by using a trellis decoding (cf. C. Yuan, unpublished results, with permission).

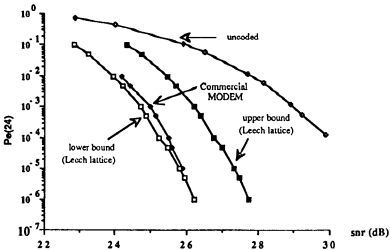


Figure 6. Block error rate vs signal-to-noise ratio for the Leech lattice.

The data labelled "commercial modem" concern a 19,200 bits/s modem (7 (bits/s)/Hz, 2,742.86 bauds) manufactured by Motorola, according to [18]

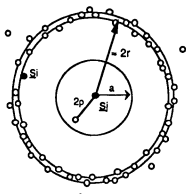


Figure 7. Pictorial representation of the signal constellation in Euclidean space \mathbb{R}^n , n very large.

Eyesight from a particular signal s_j . Almost all other signal points lie on a hypersphere of radius close to $2a = 2r$. The picture does not change if the view point becomes the one which represents another signal, for instance s_i .

CHAPITRE IV

THEORIE DE LA MODULATION CODEE ALEATOIRE

IV.1. PERFORMANCE DU CODAGE-MODULATION POUR LE CANAL GAUSSIEN

Récemment, Battail a montré que les systèmes utilisant un codage du type MDS sur un alphabet de grand taille conduisent à une distribution des distances voisine de celle que l'on obtient en moyenne par codage aléatoire [BAT *et al.* 1990, BAT 1991]. Ainsi, les performances de ces systèmes peuvent être estimées selon la technique proposée par Shannon [SHA 1949]. On a donc entrepris dans cet esprit une analyse des systèmes combinant codage et modulation, pour le canal gaussien sans ou avec fluctuations. Ce dernier modèle s'avère fréquemment intéressant, par exemple, dans le cas de liaisons radioélectriques perturbées par des phénomènes de trajet multiple et de fluctuations de niveau liés à la propagation ionosphérique en ondes décimétriques, aux masques et aux réflecteurs pour les ondes métriques et décimétriques en milieu urbain.

Nous nous sommes intéressé en particulier à l'analyse des performances du codage aléatoire sur des canaux avec évanouissement de Rayleigh. Dans ce cadre, le raisonnement utilisé dans [deO&BAT 1990] a été repris pour étudier le comportement asymptotique du taux d'erreur de ces systèmes. En présence d'évanouissement, la performance est évaluée en calculant la moyenne de la probabilité d'erreur pour le rapport signal à bruit devenu aléatoire. On a alors pu adapter les théorèmes sur le durcissement des hypersphères [p.ex. WOZ&JAC 1965] de façon à démontrer que la performance du code aléatoire sur un canal sans mémoire avec fluctuation tend vers celle d'un système sans évanouissement ayant le même rapport signal à bruit moyen. Comme corollaire, la capacité d'un canal avec évanouissement (en supposant l'entrelacement parfait) est identique à celle du canal gaussien de même rapport signal à bruit moyen [GAL 1968].

ABSTRACT

Looking for systems which combine coding and multilevel modulation whose Euclidean distance is close to that which results in the average from random coding, we consider the combination of an MDS code over a large-size alphabet and a one-to-one mapping of the alphabet into a symmetric constellation e.g., phase modulation. Its performance in the presence of additive Gaussian noise can be predicted from that of random coding, provided the signal-to-noise ratio is small enough. The results exhibit the sphere hardening phenomenon whether or not amplitude fluctuations are present. Weighted demodulator output and soft decoding should be effected in order to achieve this performance. Such a decoding can be done in principle according to previous works by Fang and Battail. A prohibitive complexity can be avoided only at the expense of strict optimality.

IV.1.1. INTRODUCTION

Battail was recently led to question the relevance of the minimum distance criterion when applied to long block codes [BAT 1989]. As a better criterion, he proposed a proximity measure of the code distance distribution with respect to that which results in the average from random coding. The criticism of the minimum distance criterion, as well as the proposed one, were formulated in the case of the Hamming distance, but they are also valid to the Euclidean metric.

Systems which combine coding and multilevel modulation, e.g., trellis-coded modulations to be used over the narrow-band Gaussian channel [UNG 1982], were designed for the Euclidean metric which is natural for this channel but they used the minimum distance criterion. We found it interesting to consider coded-modulation systems whose Euclidean distance distribution is good according to our own criterion.

Looking for such a system, we first notice that the Hamming distance distribution of the maximum-distance-separable codes (MDS), a class which especially includes the Reed-Solomon codes, is close to that which results in the average from random coding if the alphabet size and the code dimension are large [CHE 89]. In order to define a coded-modulation system using these codes, we must define the mapping of the code symbols into a constellation.

Let us consider a system using an MDS code over a large-size alphabet, associated with a symmetric constellation of the same size. We are specially interested in the case of phase modulation because of its robustness against non-linearities and amplitude fluctuations. Let first the mapping of the code symbols into this constellation be random and vary according to the position in the word.

The distance distribution of MDS codes and the random character of the mapping of the code symbols result in a distribution of Euclidean distances close to that obtained using random coding and nonrandom mapping, provided the alphabet size and the code dimension are large enough. Furthermore, the Euclidean distance distribution is not actually modified if the mapping from the code alphabet to the constellation is nonrandom and does not vary with the symbol position. Performance close to that of random coding may thus be expected from this fully deterministic scheme (we already arrived at a similar conclusion via a different approach [deO&BAT 1990], but the coded-modulation system here is completely explicit). Decoding is much less complex, however, than if random coding were used. Moreover, in the case of phase modulation, this scheme easily achieves rotational invariance in order to cope with an arbitrary phase reference, unlike trellis-coded modulation.

IV.1.2. A POWERFUL TOOL: THE RANDOM CODING

Coding here is defined just by taking at random M points in an n dimensional Euclidean space which have uniform density probability on the surface of the hypersphere to which it is confined by the power limitation. The transmission rate per dimension is then $R = 1/n \ln(M)$ nats. The average error probability of such schemes on the additive Gaussian noise channel is, according to Shannon [SHA 1949]:

$$P_e = 1 + \int_0^\pi [1 - \omega_n(\theta)]^{M-1} dQ_n(\theta), \quad (1)$$

where

n is the number of dimensions,

$Q_n(\theta)$ is defined as the probability that the point which represents the received signal is moved by the noise outside a cone with vertex at the origin, whose axis is colinear with the transmitted vector and whose half-angle is denoted by θ , and

$\omega_n(\theta)$ is the normalized solid angle of such a cone in the Euclidean space. In terms of the solid angle of this cone ^(*), Ω_n , it is given by

$$\omega_n(\theta) := \Omega_n(\theta)/\Omega_n(\pi), \quad (2)$$

where

$$\Omega_n(\theta) = C(n) \int_0^\theta \sin^{n-2} x \, dx, \quad (3)$$

and $C(n)$ is an irrelevant constant which does not depend on θ .

The Riemman-Stieltjes integral (1) can be rewritten as a product of two factors depending on the dimensionality:

$$P_e = 1 - \int_0^\pi F_n(R; \theta) G_n(S/N; \theta) d\theta, \quad (4)$$

where

a) $F_n(R; \theta) := [1 - \omega_n(\theta)]^{M-1}$ which is a function of the information rate per dimension (nats/dim),

b) $G_n(S/N; \theta) := -Q'_n(\theta)$ which depends on the signal-to-noise ratio (we notice that it is positive).

(*) The solid angle of a region on the surface of an n -sphere is the ratio of the area of this region to the $(n-1)$ -th power of the radius of the hypersphere.

Examining each of these two factors, Battail has shown [BAT *et al.* 1990] that for n increasing without limit,

$$\lim_{n \rightarrow \infty} F_n(R; \theta) = \begin{cases} 1 & \theta < \theta_R := \arcsin(\exp(-R)) \\ 0 & \text{otherwise,} \end{cases} \quad (5a)$$

and

$$\lim_{n \rightarrow \infty} G_n(S/N; \theta) = \delta(\theta - \theta_C), \quad (5b)$$

where δ is a Dirac distribution [BOU 1964] and $\theta_C := \arcsin((1 + S/N))^{-1/2}$.

Therefore, the average error probability goes asymptotically to 0 or 1 according to

$$\begin{cases} 0 & \text{if } \theta_R > \theta_C, \\ 1 & \text{if } \theta_R < \theta_C. \end{cases} \quad (6)$$

If $\theta_R > \theta_C$ (which obviously implies $R < C$) then the error probability vanishes. In contrast, it approaches unity if $\theta_R < \theta_C$ (or $R > C$).

IV.1.3. PERFORMANCE PREDICTION USING RANDOM CODING

Approximations for the First Factor (a).

The exact expressions for the normalized solid angle of the cone are:

$$\omega_n(\theta) = \frac{1}{\pi} \left[\theta + 2 (-1)^{n/2-1} \sum_{i=0}^{n/2-2} (-1)^i \frac{C_{n-2}^i}{C_{n-2}^{n/2-1}} \frac{\sin[(n-2i-2)\theta]}{n-2i-2} \right], \quad n \text{ even,} \quad (7a)$$

and

$$\omega_n(\theta) = \frac{1}{2} \left[1 - \frac{\sum_{i=0}^{(n-3)/2} (-1)^i C_{n-2}^i \frac{\cos[(n-2i-2)\theta]}{n-2i-2}}{\sum_{i=0}^{(n-3)/2} (-1)^i C_{n-2}^i \frac{1}{n-2i-2}} \right], \quad n \text{ odd.} \quad (7b)$$

Interesting approximations are obtained letting $\theta = \arcsin(x)$, and considering $\omega_n(\theta) = \omega_n((x))$. Taking the first order development of the McLaurin series, we obtain

$$F_n(R; \theta) \approx \exp \left[\frac{-\exp[n(R + \ln(\sin\theta))]}{\sin\theta \sqrt{2\pi n}} \right] \quad (8)$$

F_n is a continuous function of x which tends to the unit step at the point $x = \exp(-R)$

as n increases indefinitely. The transition from 0 to 1 is the more pronounced the larger is n .

For (finite) n large enough, the transition becomes

$$\theta_R(n) \approx \arcsin[(2\pi n)^{1/2(n-1)} \exp(-R(n-1)/n)]. \quad (9)$$

so that

$$\lim_{n \rightarrow \infty} \theta_R(n) := \theta_R = \arcsin(\exp(-R)), \quad (10)$$

a result already obtained by Shannon.

Approximations for the Second Factor (b).

Shannon has shown that [SHA 1959]

$$Q_n(\theta) = P(n-1, \sqrt{nS/N}, \sqrt{n-1} \cot \theta), \quad (11)$$

where $P(\dots)$ denotes the noncentral t -distribution [KOR&KOR 1968, ABRA&STE 1972], that is,

$$P(m, a, t) = \Pr \left[Z + a \leq t \sqrt{\frac{1}{m} \sum_1^m x_i^2} \right], \quad (12)$$

where Z and x_i , $1 \leq i \leq m$, are mutually independent Gaussian random variables, with zero mean and variance unity; a and t are constant parameters.

Approximations for $Q'_n(\theta)$ can be found in Battail (*ibidem*).

We give here another simple approximation for $Q_n(\theta)$, n large, which can be derived by taking a Gaussian approximation to the χ^2 -function [ABRA&STE 1972]:

$$Q_n(\theta) \approx 1 - Q \left(\frac{\sqrt{n-1} \cot \theta \left[1 - \frac{1}{4(n-1)} \right] - \sqrt{n} \rho}{\sqrt{1 + \frac{(n-1) \cot^2(\theta)}{2(n-1)}}} \right) \quad (13)$$

Further, we can take

$$Q_n(\theta) \approx \tilde{Q}_n(\theta) := 1 - Q \left(\sqrt{n} \frac{\cot \theta - \rho}{\sqrt{1 + 0.5 \cot^2 \theta}} \right), \quad (14)$$

on the assumption that $n \gg 1$ and $0 < \theta < \pi$. Numerical evaluation of $Q_n(\theta)$ is quite easy and then we carry out its derivative.

IV.1.4. PRESENCE OF RAYLEIGH FLUCTUATIONS

In the presence of fading, the performance of the random coding must be evaluated by averaging the error probability, i.e.

$$\bar{P}_e =: E_{S/N} (P_e) = 1 - E_{S/N} \left[\int_0^\pi F_n(R; \theta) G_n\left(\frac{S}{N}; \theta\right) d\theta \right], \quad (15)$$

where $E_{S/N}(\cdot)$ denotes the expected value with respect to the random variable S/N .

$$\bar{P}_e = 1 - \int_0^\infty dS/N \int_0^\pi F_n(R; \theta) G_n(S/N; \theta) P(S/N) d\theta,$$

where

$$\int_0^\infty P(S/N) dS/N = 1,$$

$$|F_n(R; \theta)| \leq 1, \text{ and } \int_0^\pi |G_n(S/N; \theta)| d\theta = 1.$$

Changing the order of the integrations (it may be done from the above), we obtain:

$$\bar{P}_e = 1 - \int_0^\pi F_n(R; \theta) E_{S/N} \left[G_n\left(\frac{S}{N}; \theta\right) \right] d\theta. \quad (16)$$

Furthermore, defining $\overline{G_n(\rho; \theta)} =: E_{S/N} [G_n(\frac{S}{N}; \theta)]$, it follows that

$$\bar{P}_e = 1 - \int_0^\pi F_n(R; \theta) \overline{G_n(\rho; \theta)} d\theta. \quad (17)$$

It can be noticed that fluctuations affect just the $G_n(\frac{S}{N}; \theta)$ term which depends on the signal-to-noise ratio.

Remembering that $G_n(\frac{S}{N}; \theta) = -\frac{d Q_n(\theta)}{d\theta}$, we find out

$$\overline{G_n(\rho; \theta)} = -\frac{d}{d\theta} E_{S/N} [Q_n(\theta)]. \quad (18)$$

The above equation can be rewritten as

$$\overline{G_n(\rho; \theta)} = -\frac{d}{d\theta} \overline{Q_n(\theta)} \quad (19)$$

where we have defined

$$\overline{Q_n(\theta)} =: E_{S/N} [Q_n(\theta)] = E_{S/N} [P(n-1, \sqrt{nS/N}, \sqrt{n-1} \cot \theta)]. \quad (20)$$

Here, we shall follow closely de Oliveira-Battail [deO&BAT 1990] in order to evaluate the asymptotic behaviour of $\overline{Q_n(\theta)}$.

In this case, we have yet $Q_n(\theta) = P(\dots)$ $n=2l$, but the parameter $\alpha^2 = nS/N$ becomes a random variable (r.v.).

An interleaver is a device that rearranges the ordering of a sequence of symbols in some one-to-one deterministic manner. An (n_1, n_2) interleaver reorders a sequence so that no contiguous sequence of n_2 symbols in the reordered sequence contains any pair of symbols that were separated by fewer than n_1 symbols in the original ordering, n_1 and n_2 being relative prime [RAM 1970].

We can design an appropriate interleaver to make the channel virtually memoryless (i.e. a vanishing correlation between symbols). Therefore, the multiplicative amplitudes associated with each symbol of a codeword are mutually independent so the corresponding probability density $P_\alpha(\cdot)$ is the convolution product of $l = n/2$ Rayleigh distributions yielding a chi-square distribution with $2l$ degrees of freedom.

Let us first consider some random variable transformations and their related probability density functions (pdf's):

$$P_{\alpha}(n; \xi) = \frac{1}{2^{l-1} \rho^l \Gamma(l)} \xi^{2l-1} \exp(-\xi^2/2\rho), \quad (21.a)$$

$$P_{\alpha^2}(n; \xi) = \frac{1}{2^l \Gamma(l)} \xi^{l-1} \exp(-\xi/2\rho), \quad (21.b)$$

$$P_{\alpha^2/n}(n; \xi) = \frac{l/\rho^l}{\Gamma(l)} \xi^{l-1} \exp(-l\xi/\rho), \quad (21.c)$$

where $P_X(\xi) = P(X \leq \xi)$ is the pdf of the r.v. X .

Indeed, ρ denotes the average signal-to-noise ratio as we can see by taking

$$E(\alpha^2) = \int_0^\infty \xi^2 P_{\alpha}(n; \xi) d\xi = \int_0^\infty \xi P_{\alpha^2}(n; \xi) d\xi = n\rho, \quad (22)$$

so that

$$E\left(n \frac{S}{N}\right) = n\rho \quad \text{or} \quad E\left(\frac{S}{N}\right) = \rho. \quad (23)$$

In order to evaluate the (average) error probability, we need to average $Q_n(\theta) = P(n-1, \alpha, \sqrt{n-1} \cot \theta)$ with respect to the random parameter α , as mentioned earlier. This can be carried out alternatively by:

$$\overline{Q_n(\theta)} = \int_0^\infty P(n-1, \xi, \sqrt{n-1} \cot \theta) P_{\alpha}(n; \xi) d\xi, \quad (24.a)$$

$$= \int_0^{\infty} P(n-1, \sqrt{\xi}, \sqrt{n-1} \cot \theta) P_{\alpha^2}(n; \xi) d\xi, \quad (24.b)$$

$$= \int_0^{\infty} P(n-1, \sqrt{n\xi}, \sqrt{n-1} \cot \theta) P_{\alpha^2/n}(n; \xi) d\xi. \quad (24.c)$$

We shall consider the following set of functions:

$$Q_{\alpha}(n; \xi, \theta) = P(n-1, \xi, \sqrt{n-1} \cot \theta), \quad (25.a)$$

$$Q_{\alpha^2}(n; \xi, \theta) = P(n-1, \sqrt{\xi}, \sqrt{n-1} \cot \theta), \quad (25.b)$$

$$Q_{\alpha^2/n}(n; \xi, \theta) = P(n-1, \sqrt{n\xi}, \sqrt{n-1} \cot \theta), \quad (25.c)$$

so $\overline{Q_n(\theta)}$ can be rewritten using the above definitions, yielding

$$\overline{Q_n(\theta)} = \int_0^{\infty} Q_{\alpha}(n; \xi, \theta) P_{\alpha}(n; \xi) d\xi = E_{\alpha}[Q_{\alpha}(n; \xi, \theta)], \quad (26.a)$$

$$= \int_0^{\infty} Q_{\alpha^2}(n; \xi, \theta) P_{\alpha^2}(n; \xi) d\xi = E_{\alpha^2}[Q_{\alpha^2}(n; \xi, \theta)], \quad (26.b)$$

$$= \int_0^{\infty} Q_{\alpha^2/n}(n; \xi, \theta) P_{\alpha^2/n}(n; \xi) d\xi = E_{\alpha^2/n}[Q_{\alpha^2/n}(n; \xi, \theta)]. \quad (26.c)$$

Again, $E_X[\cdot]$ denotes the expected value with respect to the r.v. X .

We will need the two following theorems (see [deO&BAT 1990] for proof):

THEOREM 1. (Weak-law theorem).

Let $\alpha^2 = nS/N$ be a random variable with distribution given by (21) where ρ is the average signal-to-noise ratio. Then $\text{plim} (\frac{S}{N}) = \rho$, that is, for any positive real $\epsilon > 0$ we have $\lim P(|\frac{S}{N} - \rho| \geq \epsilon) = 0$. ■

THEOREM 2.

The sequence of functions $[P_{\alpha^2/n}(n; \xi)]_1^{\infty}$ defines a Dirac function $\delta(\xi - \rho)$. ■

Now, we intend to prove that the presence of Rayleigh fading does not change the performance with respect to a constant signal-to-noise channel, provided the dimensionality increases without limit, a known result [GAL 1968].

THEOREM 3. The performance of the random code on a memoryless Rayleigh fading channel tends, as dimensionality increases, to the performance of a system with no fading and the same mean signal-to-noise ratio. ■

proof. Taking the limit when n increases indefinitely in (26.c), we obtain

$$\lim_{n \rightarrow \infty} \overline{Q_n(\theta)} = \lim_{n \rightarrow \infty} E_{\alpha^2/n} [Q_{\alpha^2/n}(n; \xi, \theta)]. \quad (27)$$

On the other hand, it follows from theorem 2 that $\lim_{n \rightarrow \infty} P_{\alpha^2/n}(n; \xi) = \delta(\xi - \rho)$, so that

$$\lim_{n \rightarrow \infty} \overline{Q_n(\theta)} = Q_{\alpha^2/n}(n; \rho, \theta) = Q_n(\theta) \Big|_{\frac{S}{N} = \rho}. \quad (28)$$

Applying this in (19), we have

$$\lim_{n \rightarrow \infty} \overline{G_n(\theta)} = G_n\left(\frac{S}{N}, \theta\right) \Big|_{\frac{S}{N} = \rho}. \quad (29)$$

Finally, substituting (29) in (17) it follows the claimed result, $\overline{P_e} = P_e \Big|_{\frac{S}{N} = \rho}$.

Q.E.D.

COROLLARY.

The capacity of a memoryless Rayleigh-fading channel is the same as that of a Gaussian channel with the same mean signal-to-noise ratio. ■

The numerical results obtained closely agree with theorem 3.

IV.1.5. CONCLUSION

It is known that MDS codes are good. It is our opinion that they are still better than currently believed and that, when combined with a proper multilevel modulation, they can behave very similarly to random coding, thus closely approaching the capacity of the channel. The key to their efficient use is weighted decoding since it can approach maximum likelihood decoding which chooses the signal point closest to the received one according to the Euclidean metric. Complexity can be kept reasonably small only at the expense of strict optimality.

IV.2. THE RANDOM CODED MODULATION: PERFORMANCE AND EUCLIDEAN DISTANCE SPECTRUM EVALUATION.

LE MOT MAGIQUE

*Dans un espace à très grande dimension
habitent des milliers de bits d'information
à des années-lumières de distance euclidienne
noyés par le bruit d'une mer gaussienne
des petits points tels qu'étoiles brillantes
transmis à un rythme hallucinant
(quoique contraint par la capacité rayonnante)
exhibent des taux d'erreur extravagants
mais bornés par une exponentielle dominante ...
O univers de constellations omnipotentes!
un peu partout, des voix quasi omniscientes
tantôt crient tantôt murmurent, redondantes,
un mot magique, le nom "Shannon" ...*

Deux concepts ont principalement été exploités dans ce chapitre: le raisonnement de Shannon utilisant le codage aléatoire et la critique du critère de distance minimale revendiquée par Battail. J'ai proposé d'appliquer à la modulation codée le célèbre raisonnement du codage aléatoire. On considère d'abord une constellation multidimensionnelle non codée créée par la concatenation de $N/2$ constellations bidimensionnelles, c'est-à-dire $N/2$ fois le produit cartésien d'une constellation constituante par elle-même. Alors, une constellation aléatoire est construite en choisissant au hasard un certain nombre de signaux, inférieur au nombre total de signaux disponibles, dans la constellation multidimensionnelle non codée. Une borne simple sur la probabilité moyenne d'erreur a été obtenue par la borne de la réunion. L'idée d'un débit critique partiel (en fonction de la dimensionnalité disponible) a été développée. Ensuite, j'ai introduit le concept de série θ pour les constellations bornées qui généralise les séries θ connues pour les réseaux. J'ai développé des bornes de la probabilité d'erreur par bloc et par symbole, basées sur la fonction de transfert du jeu de signaux employé. Ces bornes ont une forte analogie avec les bornes de la probabilité d'erreur des codes convolutifs introduites par Viterbi [VIT 1971]. J'ai ensuite étudié les caractéristiques de la distribution euclidienne des distances. L'évaluation de paramètres tels que la moyenne et la variance du carré de la distance a encore pu être faite à l'aide de la fonction de transfert. Les spectres de distance de quelques constellations multidimensionnelles aléatoires ont été obtenues dans le cas de constellations constituantes carrées ou en croix. J'ai mis en évidence un phénomène assez

intéressant (ce que j'appelle le durcissement de la distance normalisée) de concentration des valeurs du carré de la distance euclidienne autour de sa moyenne. Cela montre que la distance minimale n'est pas importante dans le cas d'un grand nombre de dimensions, mais plutôt la distance moyenne. Les critiques faites dans le cas du codage classique [BAT 1989] se généralisent donc aussi dans le cas de la modulation codée. Ces résultats rejoignent clairement l'idée de Battail. Enfin, j'ai démontré de même que l'énoncé "tous les codes sont bons" est vérifié aussi pour la modulation codée. En fait, excepté pour un ensemble de mesure nulle, toutes les constellations présentent une probabilité d'erreur aussi petite que l'on veut quand le nombre de dimensions de la constellation codée tend vers l'infini.

ABSTRACT

We apply the random coding argument to coded modulation. The well-known union bound on the error probability of general signaling schemes is revisited. The random coded modulation idea is introduced and a simple bound on the average performance of coded constellations is presented. A relationship between the union bound and the cut-off rate is exhibited by introducing the concept of N -dimensional partial cut-off rate. We define finite theta series for bounded finite-dimensional constellations and their related transfer functions. Bounds on the block and symbol error probability are developed which are based on the transfer function. The discussion is then focused on the squared Euclidean distance distribution. The evaluation of such parameters as its first two moments (average squared distance and squared distance variance) is considered by either finite theta series or transfer function of the bounded signal set. The Euclidean distance spectra of a few multidimensional coded modulation schemes based on square/cross constituent two dimensional constellations are presented. Their respective partial cut-off rates are computed. Finally, we discuss the asymptotic behaviour of very long coded constellations.

IV.2.1. INTRODUCTION

ONE of the deepest ideas of coding theory is the random coding argument introduced by Shannon [1]. On the other hand, many efficient coding techniques have been developed much later using the famous idea of coded modulation by Ungerboeck [2]. In this paper we intend to amalgamate these two concepts: random coding is combined with 2-dimensional modulation e.g., Quadrature Amplitude Modulation (QAM), in order to investigate the average performance of multidimensional coded constellations. We were in part motivated by the fact that some extremely regular codes appear to mimic the random code [3,4].

An upper bound on the average error probability of random coded modulation is presented in section IV.2.2. It derives from the union bound and leads to the definition of a partial cut-off rate, where "partial" means here that this quantity does take into account the actual number of dimensions. Following, we define finite theta series for bounded

multidimensional constellations and their related transfer functions. An interesting bound on the N -dimensional error probability of coded constellations is developed which is based on the transfer function and can also be applied to lattices. These bounds present a strong likeness with those on performance of convolutional codes earlier introduced by Viterbi [5], and they are also related to those on the trellis codes performance by Zehavi-Wolf [6].

In section IV.2.3 the distance spectra of multidimensional random coded constellations are investigated and examples are given. Parameters like the average squared Euclidean distance and squared distance variance of the distance spectrum are evaluated in terms of either the finite theta series or the transfer function of the constellation. The distance distribution is calculated for some multidimensional constellations based on square/cross constituent 2-dimensional (2D) constellations. Further, we prove that an interesting phenomenon occurs, namely a normalized squared distance hardening, which shows that the average squared distance becomes dominant on the error rate for high-dimensional constellations. This result confirms the fact that minimal distance is not significant for long codes [7-8]. It may also be related to our recent finding that lattices exist which can achieve the channel capacity [9]. We develop in section IV.2.4 upper bounds on the symbol error probability still based on the transfer function of the signal set.

Finally, the behaviour of very high dimensional constellations is investigated in section IV.2.5 and we prove that almost all coded constellations become good as dimensionality increases. Besides, we show that almost all coded signal sets tend to become quasi-identical in a sense to be specified later. Section IV.2.6 provides a summary and concluding remarks.

IV.2.2. BOUNDS ON THE ERROR PROBABILITY AND THE PARTIAL CUT-OFF RATE

The noise will be considered to be additive white and Gaussian (AWGN) with a double-sided spectral power density $\sigma^2 = N_0/2$ (Watts/Hz)/dim. Also, we adopt the symbol $:=$ to denote equals by definition and $\exp_2(x)$ will denote 2^x .

THEOREM 1 (UNION BOUND): The error probability $P_e(N)$ of an N -dimensional constellation with M equally likely signals over an AWGN channel is upper bounded by

$$P_e(N) \leq \frac{1}{2} \sum_{d=d_{\min}}^{d_{\max}} N_{\text{av}}(d^2) \operatorname{erfc}\left(\sqrt{\frac{d^2/4}{P(C)}\gamma_{\text{av}}}\right)$$

where $\gamma_{\text{av}} = E_{\text{av}}/N_0$ is the signal-to-noise ratio,

$P(C)$ is the average normalized power per 2-dimensional symbol,

$N_{\text{av}}(d^2)$ is the average number of signal pairs with squared Euclidean distance d^2 .

proof. Let $\{\underline{s}_0, \underline{s}_1, \dots, \underline{s}_{M-1}\}$ denote the signals of a (coded or uncoded) N -dimensional constellation. Since \underline{s}_i is a point in an N -dimensional space, it can be written in terms of its coordinates, that is, $\underline{s}_i = (s_{i1}, s_{i2}, \dots, s_{iN-1}, s_{iN})$. The squared Euclidean

distance between two signal points, \underline{s}_i and \underline{s}_j , is given by

$$d^2(\underline{s}_i, \underline{s}_j) = \sum_{k=1}^N d^2(s_{ik}, s_{jk}). \quad (1)$$

On the other hand, the error probability $P_e(N)$ of an N -dimensional signal set can be evaluated by

$$P_e(N) = \sum_{i=0}^{M-1} P(\underline{s}_i) P(e | \underline{s}_i) = \frac{1}{M} \sum_{i=0}^{M-1} P(e | \underline{s}_i) \quad (2)$$

where $P(\underline{s}_i)$ is the probability that \underline{s}_i is transmitted, and

$P(e | \underline{s}_i)$ is the conditional error probability.

In order to calculate the conditional error probability given that the signal \underline{s}_i was transmitted, we shall consider an index set $T_i := \{j \in 0, 1, \dots, M-1 \mid j \neq i\}$. Certainly, an error is made iff the decoded signal \underline{s} belongs to the error set $E_i := \bigcup_j \underline{s}_j \mid j \neq i$.

Hence,

$$P(e | \underline{s}_i) = P(\text{dec } \underline{s} \in E_i) = P(\text{dec } \underline{s} \in \bigcup_{j \in T_i} \underline{s}_j) \quad (3)$$

where $\{\text{dec } \underline{s}\}$ denotes the event "decoding the signal \underline{s} ". For simplicity, we dropped out "given that \underline{s}_i was transmitted".

We can make a partition of the index set T_i according to the distance between \underline{s}_i and the elements whose index belongs to the set T_i . Let $T_i^d := \{j \in 0, 1, 2, \dots, M-1 \mid d^2(\underline{s}_i, \underline{s}_j) = d^2\}$ be such a partition of T_i . Then $T_i^d \subset T_i$ and $\bigcup_{d=d_{\min}} T_i^d = T_i$. Furthermore, $T_i^d \cap T_i^{d'} = \emptyset$ for $d \neq d'$, \emptyset being the empty set. This partition over the index set induces a partition over E_i which is given by the following subsets:

$$E_i^d := \bigcup_j \{\underline{s}_j \mid d^2(\underline{s}_i, \underline{s}_j) = d^2\} = \bigcup_{j \in T_i^d} \underline{s}_j \quad (4)$$

in such a way that $E_i^d \cap E_i^{d'} = \emptyset$ for $d \neq d'$.

Applying these partitions, the error set E_i can be rewritten as $E_i = \bigcup_{d=d_{\min}}^{d_{\max}} \bigcup_{j \in T_i^d} \underline{s}_j$.

Consequently, it follows from (3) that

$$P(e | \underline{s}_i) = P(\text{dec } \underline{s} \in \bigcup_{d=d_{\min}}^{d_{\max}} \bigcup_{j \in T_i^d} \underline{s}_j \rightarrow \underline{s}_j) \quad (5)$$

where $\underline{s}_i \rightarrow \underline{s}_j$ denotes the event of decoding \underline{s}_j when \underline{s}_i is sent and \underline{s}_j is considered as the only alternative. The use of the union bound (see e.g. [18], p.60) to the above equation yields:

$$P(e | \underline{s}_i) \leq \sum_{d=d_{\min}}^{d_{\max}} \sum_{j \in T_i^d} P(\underline{s}_i \rightarrow \underline{s}_j) \quad (6)$$

where $P(\underline{s}_i \rightarrow \underline{s}_j)$ denotes the pairwise error probability when \underline{s}_i is sent and \underline{s}_j is the only alternative. It should be pointed out that $P(\underline{s}_i \rightarrow \underline{s}_j)$ for $j \in T_i^d$ does not depend on j but just on d^2 . Moreover, it can easily be shown that

$$P(\underline{s}_i \rightarrow \underline{s}_j) = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{(d/2)^2}{2\sigma^2}}\right) = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{d^2/4}{P(C)}} \gamma_{\text{av}}\right) \quad (7)$$

where the second equality results from the equality $\gamma_{\text{av}} = E_{\text{av}}/N_0 = P(C)/2\sigma^2$. To simplify notation, the pairwise error probability of any two points i, j with square distance d^2 will be denoted throughout by $e(d^2) := \frac{1}{2} \operatorname{erfc}(\sqrt{ad^2})$, where $a := \gamma_{\text{av}}/4P(C)$.

Now, let $N_i(d^2) = |T_i^d|$ denote the cardinality of the set T_i^d , i.e., the number of signals at a squared Euclidean distance d^2 from \underline{s}_i . It results from equalities (6) and (7) that

$$P(e | \underline{s}_i) \leq \sum_{d=d_{\min}}^{d_{\max}} N_i(d^2) e(d^2). \quad (8)$$

Therefore, inserting into (2) we find

$$P_e(N) \leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{d=d_{\min}}^{d_{\max}} N_i(d^2) e(d^2). \quad (9)$$

Finally, the proof is completed by changing the order of the summations and defining

$$N_{\text{av}}(d^2) := \frac{1}{M} \sum_{i=0}^{M-1} N_i(d^2) \quad (10)$$

which represents the average number of points at a distance d^2 in the signal constellation. Q.E.D.

Indeed, this theorem applies to both coded and uncoded constellations. As a simple example, we consider the uncoded M -ary 2-dimensional constellation corresponding to narrow-sense QAM. Assuming a high signal-to-noise ratio ($\frac{d_{\min}^2/4}{P(C)} \gamma_{\text{av}} \gg 1$) and remembering that $d_{\min} = 2$ and $N_{\text{av}}(d_{\min}^2) = 4$ (for large M), we derive the well-known approximation of the symbol error probability:

$$P_e(2) \approx \frac{1}{2} N_{\text{av}}(d_{\min}^2) \operatorname{erfc}\left(\sqrt{\frac{d_{\min}^2/4}{P(C)}} \gamma_{\text{av}}\right) \approx 2 \operatorname{erfc}\left(\sqrt{\frac{3\gamma_{\text{av}}}{2(M-1)}}\right). \quad (11)$$

In the following, we apply the famous random coding argument to coded modulation systems.

Random coding

Let us consider, to begin with, an N -dimensional uncoded constellation (N even) created by concatenating $N/2$ identical 2-dimensional constellations whose alphabet size is q , i.e., by taking the Cartesian product of a constituent 2-dimensional constellation with itself $N/2$ times. From its total number of uncoded signals $M_{\text{unc}} = q^{N/2}$, we build an N -

dimensional random coded constellation by a random (uniform) choice of $M_{\text{cod}} = q^{K/2}$ points, ($M_{\text{cod}} < M_{\text{unc}}$), resulting in a rate of

$$R = \frac{1}{N} \log_2 M_{\text{cod}} = \frac{1}{2} \frac{K}{N} \log_2 q \quad \text{bits/dim} \quad (12)$$

or a normalized rate [10,11] of $R = (K/N) \log_2 q$ bits per 2-dimensional symbol.

The way of introducing redundancy by increasing the alphabet size is the same as in Ungerboeck's work. This coding technique could be seen as a Block Code Modulation (BCM) where q -ary (N, K) block codes are picked at random. Additionally, it should be pointed out that this process deals with the ensemble of all codes and that no linearity assumption is made. By an abuse of language, a coded constellation is referred indistinctly to as a code. Henceforth, we denote by Ω an N -dimensional constellation while its constituent 2-dimensional constellation is denoted by Ω_2 . Also, the minimal distance in the constituent 2-dimensional constellation is denoted by d_0 .

Now the idea of random coded modulation is used in connection with Theorem 1 (union bound) in order to estimate the average error probability of the set of all codes. This is made according to the following theorem:

THEOREM 2. (RANDOM CODING): Let $\bar{P}_e(N)$ be the average error probability of a random coded modulation with an N -dimensional signal set. Then the following inequality holds:

$$\bar{P}_e(N) \leq \frac{1}{2} \sum_{d=d_{\min}} \bar{N}_{av}(d^2) \operatorname{erfc} \left(\sqrt{\frac{d^2/4}{P(C)} \gamma_{\text{av}}} \right)$$

where $\bar{N}_{av}(d^2)$ denote an average unnormalized squared distance profile.

proof. We deal with L codes chosen by random means as suggested earlier. Applying the preceding theorem to each code, the following set of inequalities results:

$$P_e^{\#1}(N) \leq \sum_{d=d_{\min}} N_{av}^{\#1}(d^2) e(d^2) \quad (13.1)$$

$$P_e^{\#2}(N) \leq \sum_{d=d_{\min}} N_{av}^{\#2}(d^2) e(d^2) \quad (13.2)$$

...

$$P_e^{\#L}(N) \leq \sum_{d=d_{\min}} N_{av}^{\#L}(d^2) e(d^2) . \quad (13.L)$$

the superscript $\#l$ standing for the l -th code. We note that these codes have not the same minimal distance. We use here the minimum of the minimal distance among all conceivable codes, namely d_0 . Furthermore, there are at most $M_{\text{cod}}(M_{\text{cod}}-1)$ terms in each sum. Thus, the mean error probability of random codes can be expressed as

$$\bar{P}_e(N) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L P_e^{\#l}(N) . \quad (14)$$

Thus, it results from equalities (13.*) that

$$\bar{P}_e(N) \leq \sum_{d=d_{\min}} \bar{N}_{av}(d^2) e(d^2) \quad (15)$$

where

$$\bar{N}_{av}(d^2) := \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L N_{av}^{(l)}(d^2) \quad (16)$$

is referred to as the unnormalized squared distance profile of the random code. Q.E.D.

It should be stressed that $\bar{N}_{av}(d^2)$ is not the squared distance profile of any code of the random ensemble but rather the average profile resulting from taking into account all codes. Strictly speaking, there is no code among all existing codes whose profile is $\bar{N}_{av}(d^2)$. Now we introduce a normalized squared distance profile referred to as 'squared distance distribution'.

DEFINITION 1. The squared distance distribution (*distance spectrum*) of the uncoded and random coded constellations are respectively

$$P(d^2)_{\text{unc}} := N_{av}(d^2) / M_{\text{unc}} \quad \text{and} \quad \bar{P}(d^2) := \bar{N}_{av}(d^2) / M_{\text{cod}}$$

$N_{av}(d^2)$ standing for the distance profile of the uncoded N -dimensional constellation.

Evidently, one has $\sum P(d^2) = 1$ in both cases. The way to evaluate $\bar{N}_{av}(d^2)$, and consequently to estimate the average performance, is quite simple.

ASSUMPTION 1. We consider that the N -dimensional random coded constellation has the same distribution as the uncoded N -dimensional signal set, that is to say, $\bar{P}(d^2) = P(d^2)_{\text{unc}}$.

An immediate consequence of this assumption is

$$\bar{N}_{av}(d^2) = q^{-(N-K)/2} N_{av}(d^2) = \exp_2 \{-N [\frac{1}{2} \log_2 q - R]\} N_{av}(d^2). \quad (17)$$

We intend now to show that this union bound is closely linked to a cut-off rate \bar{R}_0 so that these bounds are useless both near and above \bar{R}_0 .

THEOREM 3. (PARTIAL CUT-OFF RATE): The average error probability of an N -dimensional random coded signal set is exponentially upper bounded by

$$\bar{P}_e(N) \leq \exp_2 \{-N [\bar{R}_0(N) - R]\}$$

where $\bar{R}_0(N) = -1/N \log_2 E[e(d^2)]$ is the N -dimensional partial cut-off rate.

proof. Of course, the average error probability is upper bounded according to Theorem 2 which can be rewritten as

$$\bar{P}_e(N) \leq \exp_2 \left\{ -N \left[-\frac{1}{N} \log_2 \left(\sum_{d=d_{\min}} \bar{N}_{av}(d^2) e(d^2) \right) \right] \right\}. \quad (18)$$

Besides, a reliability function, namely $E(R, \bar{R}_0)$, is defined as

$$E(R, \bar{R}_0) := -1/N \log_2 \sum_{d=d_{\min}} \bar{N}_{av}(d^2) e(d^2) \quad (19)$$

so equation (18) can be expressed in the form

$$\bar{P}_e(N) \leq \exp_2 \left[-N E(R, \bar{R}_0) \right]. \quad (20)$$

Remembering that $M_{\text{cod}} = \exp_2(NR)$, R in bits/dimension, we put

$$E(R, \bar{R}_0) = -\frac{1}{N} \log_2 \left\{ \exp_2(NR) \sum_{d=d_{\min}} \frac{\bar{N}_{av}(d^2)}{M_{\text{cod}}} e(d^2) \right\} \quad (21)$$

yielding

$$E(R, \bar{R}_0) = [-1/N \log_2 \sum_{d=d_{\min}} \bar{P}(d^2) e(d^2)] - R. \quad (22)$$

We now define the N -dimensional partial cut-off rate as

$$\bar{R}_0(N) := -\frac{1}{N} \log_2 E \left[\frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{d^2/4}{P(C)}} \gamma_{av} \right) \right] \quad \text{bits/dim.} \quad (23)$$

where the expectation is taken with respect to the distribution $\bar{P}(d^2)$, the term $\bar{P}(0)$ not being taken into account. This leads to $E(R, \bar{R}_0) = \bar{R}_0(N) - R$ and the proof follows. Q.E.D.

DEFINITION 2. The cut-off rate, \bar{R}_0 , is related to $\bar{R}_0(N)$ by $\lim_{N \rightarrow \infty} \bar{R}_0(N) := \bar{R}_0$, where both $\bar{R}_0(N)$ and \bar{R}_0 are expressed in bits per dimension.

It should be pointed out that Theorem 3 can be used for a particular code (coded constellation) as well as for the random code: it suffices to apply Theorem 1 instead of Theorem 2 in its derivation. Hence, an interesting way to compare different codes is to calculate their respective partial cut-off rates, i.e. for the l^{th} code:

$$R_0^{\#l}(N) := -\frac{1}{N} \log_2 \sum_{d=d_{\min}} P^{\#l}(d^2) e(d^2). \quad (24)$$

The evaluation of such partial rates involves the knowledge of the complete distance spectrum which is, to some extent, hard to obtain. Hence, we consider two special rates related to the partial cut-off rate:

$$\bar{R}_0(N)_m := -\frac{1}{N} \log_2 e(d_m^2); \quad (25)$$

the subscript m standing here for min or mean. In the latter case we define $d_{\text{mean}} := [E(d^2)]^{1/2}$. The rate $\bar{R}_0(N)$ can be estimated by the following theorem:

THEOREM 4. For any dimensionality N , the partial cut-off rate verifies the following bounds: $\bar{R}_0(N)_{\min} \leq \bar{R}_0(N) \leq \bar{R}_0(N)_{\max}$

Sketch of the proof. To prove this, observe that the first bound is an immediate consequence of $\text{erfc}(\cdot)$ being monotonically decreasing with d^2 and that the second one can be proved by Jensen's inequality. **Q.E.D.**

When N is large enough, the asymptotic behaviour of the $\bar{R}_0(N)$ bounds expressed in nats/dimension is given by

$$\bar{R}_0(N)_{\min} \sim \frac{d_{\min}^2}{N} \frac{\gamma_{\text{av}}}{4P(C)} \quad \text{and} \quad \bar{R}_0(N)_{\max} \sim \frac{d_{\max}^2}{N} \frac{\gamma_{\text{av}}}{4P(C)}. \quad (26)$$

In order to compare the performance of uncoded and coded constellations, we must use the same criterion for measuring the error rate. The probability of error of uncoded constellations are currently expressed per 2-dimensional symbol, so the following lemma could be useful.

LEMMA 5. The 2-dimensional error probability of symmetrical constellations (like lattice or randomly coded ones) is related to the N -dimensional error probability by $\frac{2}{N} P_e(N) \leq P_e(2) \leq P_e(N)$.

proof. The N -dimensional error rate is expressed as

$$P_e(N) = P\left(\bigcup_{i=1}^{N/2} \text{error in the } i\text{-th 2D symbol}\right) \leq \sum_{i=1}^{N/2} P_e(2) = \frac{N}{2} P_e(2), \quad (27)$$

where we have used both the union bound and the constellation symmetry. Furthermore, the 2-dimensional error probability can be evaluated by

$$P_e(2) = \frac{E[\text{number of errors per ND}]}{N/2}, \quad (28)$$

the word "error" standing here for "erroneous 2D symbol". The average number of errors can be expressed as

$$E[\# \text{ errors per ND}] = \sum_{i=0}^{N/2} i P(i \text{ errors per ND}) \leq \frac{N}{2} \sum_{i=1}^{N/2} P(i \text{ errors per ND}). \quad (29)$$

Additionally,

$$P_e(N) = P\left(\bigcup_{i=1}^{N/2} i \text{ errors per ND}\right) = \sum_{i=1}^{N/2} P(i \text{ errors per ND}). \quad (30)$$

The proof is completed by combining (29) and (30). **Q.E.D.**

It has long been recognized that good multidimensional coded constellations can be obtained using dense finite-dimensional lattices. A lattice code Ω is defined here as a subset of a lattice Λ whose points lie within a bounded region [12]. Clearly, the dimension of a lattice code is the same as that of the taken lattice, that is, $\dim \Omega = \dim \Lambda$. The Θ series

of an (unbounded) lattice Λ is given by [12,13]

$$\Theta_{\Lambda}(z) = \sum_{d=0}^{\infty} N(d^2) w^{d^2} = 1 + \sum_{d=d_{\min}}^{\infty} N(d^2) w^{d^2} \quad (31)$$

where $w = \exp(\pi iz)$ ($i = \sqrt{-1}$) and $N(d^2)$ is the number of lattice points at a squared distance d^2 from the origin. Further, we suggest to define a finite theta series of a bounded lattice Ω according to

DEFINITION 3. (FINITE THETA SERIES):

$$\Theta_{\Omega}(z) := \sum_{d=0}^{d_{\max}} N_{\text{av}}(d^2) w^{d^2} = 1 + \sum_{d=d_{\min}}^{d_{\max}} N_{\text{av}}(d^2) w^{d^2}$$

where $w = \exp(\pi iz)$ and $N_{\text{av}}(d^2)$ denotes the average number of points at a squared distance d^2 from an arbitrary point, as defined in Theorem 1.

The total number of constellation points is $M = |\Omega|$, the cardinality of the set Ω . This function contains all the information concerning the *average* neighbourhood in the constellation. We mention that finite theta series, as defined above, can be employed not only for lattice codes but also for any bounded signal set.

PROPOSITION 6. The finite theta series verifies the properties:

$$\Theta_{\Omega}(z)|_{w=1} = \sum_{d=0}^{d_{\max}} N_{\text{av}}(d^2) = M = |\Omega| \quad \text{and} \quad \lim_{M \rightarrow \infty} \Theta_{\Omega}(z) \rightarrow \Theta_{\Lambda}(z).$$

Indeed, the bounded lattice Ω approaches the lattice Λ when M increases indefinitely. It can be seen that $N_i(d^2) \rightarrow N(d^2)$ becomes independent from i as $M \rightarrow \infty$, and that $N_{\text{av}}(d^2) \rightarrow N(d^2)$.

A bound on the performance of lattice coded systems can now be developed which is based on theta series and Theorem 1. In fact, we use transfer functions which result from the theta series (either infinite or finite) by a change in its indeterminate. A somewhat strong analogy is found with bounds on the convolutional codes performance by Viterbi [5]. This bound is attributed to Forney in [13], p.73.

LEMMA 7. The N -dimensional error rate of a coded constellation Ω taken from a lattice Λ is upper bounded by

$$P_e(N)|_{\Omega} \leq \frac{1}{2} T(D; \Omega)|_{D=D_0} \leq \frac{1}{2} T(D; \Lambda)|_{D=D_0}$$

where $T(D; \cdot)$ are the transfer functions defined below and $D_0 := \exp(-a d_0^2)$, d_0 being the minimal distance of the uncoded constituent 2D constellation.

proof. Consider the Λ -transfer function, $T(D)$ $0 < D < 1$, of a lattice Λ defined as

$$T(D; \Lambda) := \sum_{d=d_{\min}}^{d_{\max}} N(d^2) D^{d^2/d_0^2} = \Theta_{\Lambda}(z) \Big|_{w=D^{1/d_0^2}} - 1. \quad (32)$$

Similarly, the Ω -transfer function of a lattice code Ω is defined by

$$T(D; \Omega) := \sum_{d=d_{\min}}^{d_{\max}} N_{\Omega}(d^2) D^{d^2/d_0^2} = \Theta_{\Omega}(z) \Big|_{w=D^{1/d_0^2}} - 1. \quad (33)$$

On the other hand, the exponential bound on the $\operatorname{erfc}(\cdot)$ function ^(*) results in:

$$\operatorname{erfc}(\sqrt{ad^2}) \leq \exp(-ad^2) = D^{d^2/d_0^2} |_{D=D_0}. \quad (34)$$

Applying this inequality to the upper bound of Theorem 1 readily results in the (weaker) bound

$$P_e(N)|_{\Omega} \leq \frac{1}{2} \sum_{d=d_{\min}}^{d_{\max}} N_{\Omega}(d^2) D^{d^2/d_0^2} |_{D=D_0} = \frac{1}{2} T(D; \Omega) |_{D=D_0}. \quad (35)$$

Furthermore, since $N_{\Omega}(d^2) \leq N(d^2)$, it is easily verified that $(\forall D) T(D; \Omega) \leq T(D; \Lambda)$, completing the proof. **Q.E.D.**

The bound on the error probability based on the Λ -transfer function is less tight than the one based on the Ω -transfer function, but is quite straightforward to obtain. Moreover, it can be improved in the following way: a truncated transfer function $T(D; [\Lambda])$ of a lattice Λ is defined as

$$T(D; [\Lambda]) := \operatorname{TRUNC}_{\deg \Theta_{\Lambda} \geq d_{\max}} \Theta_{\Lambda}(z) \Big|_{w=D^{1/d_0^2}} - 1, \quad (36)$$

where TRUNC indicates that the Θ series written as a polynomial of w has been truncated to a finite degree but no less than d_{\max}^2 . Thus, possible convergence problems are avoided. In the cases where d_{\max} is unknown, the theta series may be truncated to the minimal degree greater than or equal to $N/2 \max_{d^2 \in S} d^2$, where $S = \Delta(\Omega_2)$ denotes the set of all conceivable squared Euclidean distances between two points in a constituent 2D constellation. As a consequence,

$$\frac{1}{2} T(D; \Omega) |_{D=D_0} \leq \frac{1}{2} T(D; [\Lambda]) |_{D=D_0} \leq \frac{1}{2} T(D; \Lambda) |_{D=D_0}. \quad (37)$$

* Of course, other bounds could be used here

IV.2.3. EUCLIDEAN DISTANCE SPECTRUM OF RANDOM CODES

Besides the minimum squared distance, we are especially interested in the first and second moments of the Euclidean distance distribution $P(d^2)$, denoted respectively by $E(d^2)$ and $E([d^2]^2)$. The next lemma presents some interesting properties relating these parameters to the theta series of a constellation.

LEMMA 8. The average and the variance of the squared Euclidean distance of a constellation Ω , respectively $E(d^2)$ and $\sigma^2(d^2)$, are given by

$$E(d^2) = \frac{1}{|\Omega|} \left. \frac{d\Theta_{\Omega}(z)}{dw} \right|_{w=1} \quad \text{and} \quad E([d^2]^2) = \frac{1}{|\Omega|} \left[\left. \frac{d^2\Theta_{\Omega}(z)}{dw^2} + \frac{d\Theta_{\Omega}(z)}{dw} \right|_{w=1} \right].$$

proof. First, we notice that

$$\left. \frac{d\Theta_{\Omega}(z)}{dw} \right|_{w=1} = \sum_{d=d_{\min}}^{d_{\max}} N_{av}(d^2) d^2 1^{d^2-1} = |\Omega| \sum_{d=d_{\min}}^{d_{\max}} P(d^2) d^2 \quad (38)$$

and the first part follows.

Furthermore,

$$\left. \frac{d^2\Theta_{\Omega}(z)}{dw^2} \right|_{w=1} = \sum_{d=d_{\min}}^{d_{\max}} N_{av}(d^2) d^2 (d^2-1) 1^{d^2-1} \quad (39)$$

which can be arranged as

$$\left. \frac{d^2\Theta_{\Omega}(z)}{dw^2} \right|_{w=1} = |\Omega| \left[\sum_{d=d_{\min}} P(d^2)(d^2)^2 - \sum_{d=d_{\min}} P(d^2)(d^2) \right] \quad (40)$$

proving the second part. **Q.E.D.**

Lemma 8 can be rewritten in terms of the transfer function of a bounded lattice as

COROLLARY 9. The first two moments of the distance distribution $P(d^2)$ are given by the formulae:

$$E(d^2) = \frac{d_0^2}{|\Omega|} \left. \frac{dT(D;\Omega)}{dD} \right|_{D=1}$$

and

$$E([d^2]^2) = \frac{d_0^4}{|\Omega|} \left[\left. \frac{d^2 T(D;\Omega)}{dD^2} \right|_{D=1} + \left. \frac{dT(D;\Omega)}{dD} \right|_{D=1} \right].$$

Besides transfer functions we are also concerned with related functions given by

DEFINITION 4. (COMPLETE TRANSFER FUNCTION): the complete transfer function, namely $T_N(D)$, of an N -dimensional lattice Λ (or an N -dimensional bounded constellation Ω) is defined as

$$T_N(D) := 1 + T(D; g) = \Theta_g(z) \Big|_{z=D^{1/d_0^2}}$$

where g stands for Λ or Ω .

It is clear heretofore if the complete transfer function concerns either a lattice or a bounded constellation.

For the multidimensional uncoded constellation specified earlier, it can be seen by a combinatory approach of the generating function [14] that

$$T_N(D) = [T_2(D)]^{N/2}. \quad (41)$$

Taking random coding into account, the complete transfer function $\bar{T}_N(D)$ of such a multidimensional constellation is given by

$$\bar{T}_N(D) = \frac{M_{\text{cod}}}{M_{\text{unc}}} T_N(D) = q^{-(N-K)/2} T_N(D) \quad (42)$$

which is an immediate consequence of Assumption 1. Also, obviously, $T_N(1) = [T_2(1)]^{N/2} = q^{N/2}$ and $\bar{T}_N(1) = q^{K/2}$, as expected (see Proposition 6).

We are interested in the squared distance distribution (spectrum) of the random coded-modulation, $\bar{P}(d^2)$, so we will use Corollary 9. Initially, we calculate the derivative of the complete transfer function. Henceforth, for the sake of simplicity, we denote $dT(D)/dD$ by $\dot{T}(D)$.

$$\dot{T}_N(D) = \frac{N}{2} [T_2(D)]^{N/2-1} \cdot \dot{T}_2(D) \quad (43.1)$$

and

$$\ddot{T}_N(D) = \frac{N}{2} \left\{ \left(\frac{N}{2} - 1 \right) [T_2(D)]^{N/2-2} \cdot [\dot{T}_2(D)]^2 + [T_2(D)]^{N/2-1} \cdot \ddot{T}_2(D) \right\}. \quad (43.2)$$

Now, we are able to establish the next result:

THEOREM 10. The mean and variance of the squared Euclidean distance between two points of a multidimensional random coded constellation are, respectively

$$E(d^2) = d_0^2 \frac{N}{2} \cdot \left\{ \frac{\dot{T}_2(1)}{T_2(1)} \right\}$$

and

$$\sigma^2(d^2) = d_0^4 \frac{N}{2} \cdot \left\{ \frac{\dot{T}_2(1)}{T_2(1)} + \frac{\ddot{T}_2(1)}{T_2(1)} - \left[\frac{\dot{T}_2(1)}{T_2(1)} \right]^2 \right\}.$$

proof. By Corollary 9, we have (the second equality results from Assumption 1):

$$E(d^2) = \frac{d_0^2}{q^{K/2}} \left. \frac{d \bar{T}_N(D)}{dD} \right|_{D=1} = \frac{d_0^2}{q^{N/2}} \left. \frac{d T_N(D)}{dD} \right|_{D=1} . \quad (44)$$

Substituting (43.1) in (44) and remembering that $T_2(1) = q$, it follows

$$E(d^2) = \frac{d_0^2}{q^{N/2}} \cdot \frac{N}{2} \cdot \frac{[T_2(1)]^{N/2}}{T_2(1)} \cdot \dot{T}_2(1) = d_0^2 \frac{N}{2} \cdot \frac{\dot{T}_2(1)}{T_2(1)} . \quad (45)$$

The second moment is given by

$$\begin{aligned} E([d^2]^2) &= \frac{d_0^4}{q^{K/2}} \left[\left. \frac{d^2 \bar{T}_N(D)}{dD} \right|_{D=1} + \left. \frac{d \bar{T}_N(D)}{dD} \right|_{D=1} \right] \\ &= \frac{d_0^4}{q^{N/2}} [\ddot{T}_N(1) + \dot{T}_N(1)] . \end{aligned} \quad (46)$$

Combining (43.1), (43.2) and the above equation, we find out after further simplifications

$$E([d^2]) = d_0^4 \frac{N}{2} \left[\frac{\dot{T}_2(1)}{T_2(1)} + \left(\frac{N}{2} - 1 \right) \frac{[\dot{T}_2(1)]^2}{[T_2(1)]^2} + \frac{\ddot{T}_2(1)}{T_2(1)} \right] . \quad (47)$$

We get the desired form of the squared distance variance by inserting (47) and (45) into $\sigma^2(d^2) = E([d^2]^2) - E^2(d^2)$, which completes the proof. **Q.E.D.**

Examples

We intend to apply the above results to a number of multidimensional random coded constellations. To begin with, let us consider a few constituent 2-dimensional constellations with alphabet size $q = 4, 8, 16, 32$ for example (figure 1). Next, we evaluate the finite theta series as well as the transfer function of such signal sets (cf. Appendix A). The respective transfer functions are presented below.

$$T_2(D) = 1 + 2D + 1D^2$$

$$T_2(D) = 1 + 2.0D + 1.5D^2 + 1.0D^4 + 2.0D^5 + 0.5D^{10}$$

$$T_2(D) = 1 + 3.0D + 2.25D^2 + 2.0D^4 + 3.0D^5 + 1.0D^8 + 1.0D^9 + 1.5D^{10} + 1.0D^{13} + 0.25D^{18}$$

$$\begin{aligned} T_2(D) = & 1 + 3.25D + 2.875D^2 + 2.5D^4 + 4.5D^5 + 1.75D^8 + 1.75D^9 + 3.25D^{10} + 2.5D^{13} + 1D^{16} + 2D^{17} + \\ & + 0.875D^{18} + 1.5D^{20} + 1.5D^{25} + 0.75D^{26} + 0.5D^{29} + 0.25D^{32} + 0.25D^{34} . \end{aligned}$$

One has $\dot{T}_2(1)/T_2(1) = 1, 3, 5, 10, \dots$ for $q = 4, 8, 16, 32, \dots$, respectively.

A simple, but striking, example of our approach makes allowance for an N -dimensional coded constellation with constituent 2-dimensional signal set of size $|\Omega_2| = 16$. Then

$$T_2(D) = 1 + 3D^1 + 2.25D^2 + 2D^4 + 3D^5 + 1D^8 + 1D^9 + 1.5D^{10} + 1D^{13} + 0.25D^{18}$$

$$\dot{T}_2(D) = 3 + 4.5D + 8D^3 + 15D^4 + 8D^7 + 9D^8 + 15D^9 + 13D^{12} + 4.5D^{17}$$

$$\ddot{T}_2(D) = 4.5 + 24D^2 + 60D^3 + 56D^6 + 72D^7 + 135D^8 + 156D^{11} + 76.5D^{16}$$

so that $T_2(1) = 16$, $\dot{T}_2(1) = 80$ and $\ddot{T}_2(1) = 584$.

Hence, assuming $d_0 = 2$, the distance spectrum is mainly characterized by: $E(d^2) = 5N/2$ and $\sigma^2(d^2) = 16.5 N/2$.

Plots of the distance spectrum for a few dimensions are exhibited in figure 2. It can be seen that their respective distributions clearly approach a Gaussian one as the dimensionality increases.

When the dimensionality increases, there is some kind of squared distance concentration near the average distance, since the squared distance spread (relative to the mean) tends asymptotically to zero. More precisely:

THEOREM 11. (NORMALIZED DISTANCE HARDENING): The squared Euclidean distance of multidimensional random coded constellations verifies

$$\lim_{N \rightarrow \infty} \frac{d^2}{N/2} = d_0^2 \frac{\dot{T}_2(1)}{T_2(1)}, \text{ where plim denotes the limit in probability.}$$

proof. We first define a variable δ , namely the 2-dimension distance (or normalized distance), according to $\delta^2 := d^2/(N/2)$. Clearly, $0 \leq \delta^2 \leq \deg T_2(D)$. It follows from Theorem 10 that

$$E(\delta^2) = d_0^2 \cdot K_1 \quad \text{and} \quad \sigma^2(\delta^2) = \frac{d_0^4 \cdot K_2}{N/2} \quad (48)$$

where K_1 and K_2 are, respectively, the two terms into brackets in Theorem 10. Thus, Chebyshev's inequality results in

$$P(|\delta^2 - E(\delta^2)| \geq \epsilon) \leq \frac{d_0^4 \cdot K_2}{N/2 \epsilon} \rightarrow 0 \quad (49)$$

which goes to zero as N increases indefinitely. **Q.E.D.**

Roughly speaking, this theorem asserts that $N(E(\delta^2)) \approx M(1 - o(1))$ for very high dimensionalities (Landau notation). Thus, there exists very few long coded constellations for which the normalized squared distance exceeds the "hardening distance" by any appreciable margin. A similar behaviour has been found for the minimal (Hamming) distance in the ensemble of binary codes having fixed ratio K/N [19], namely

$$P\left(\left|\frac{d}{N} - p\right| > \epsilon\right) \rightarrow 0 \quad \text{as} \quad N \rightarrow \infty$$

where d is the minimal Hamming distance, $H_2(p) := -p \log_2 p - (1-p) \log_2 (1-p)$ is the

binary entropy function and p is the smallest solution of $1-H_2(p) = K/N$. This asserts that very few long codes exist lying far from the Varsharmov-Gilbert bound. Now the partial cut-off rate can be written as $\bar{R}_0(N) = -1/N \log_2 E[1/2 \operatorname{erfc}(\sqrt{a\delta^2 N/2})]$.

It can be seen *ex post fact* that the average squared distance becomes dominant on the error rate performance as dimensionality increases. Thus, the minimal distance is not significant for long codes. This meets the recent criticism of the minimal distance criterion by Battail [7-8].

IV.2.4. MORE ON THE ERROR PROBABILITY BOUNDS

Naturally, the upper bounds so far introduced concern the N -dimensional (block) error probability. We are also interested in obtaining upper bounds on the random coded-modulation performance on the basis of the 2D error probability criterion (symbol error rate). To this end we introduce a second variable L in the complete transfer function, namely:

DEFINITION 5. (TWO-PARAMETER TRANSFER FUNCTION):

$$T(D, L; \Omega) := \sum_{d=d_{\min}}^{d_{\max}} \sum_{l=1}^{N/2} N_{av}(d^2, l) D^{d^2/d_0^2} L^l$$

where $N_{av}(d^2, l)$ is the average number of signal pairs with both squared distance d^2 and l different 2D symbols.

An interesting property of this new function is that $T(D, 1; \Omega) = T(D; \Omega)$, since $\sum_{l=1}^{N/2} N_{av}(d^2, l) = N_{av}(d^2)$.

THEOREM 12. The 2-dimensional error probability of a multidimensional constellation Ω is upper bounded by

$$P_e(2) \leq \frac{1}{N} \left. \frac{\partial T(D, L; \Omega)}{\partial L} \right|_{L=1, D=D_0}$$

where $T(D, L; \Omega)$ is the transfer function of Ω and D_0 as in Lemma 7.

proof. We shall follow quite closely the proof of Theorem 1. Let L be the number of erroneous symbols after decoding. The average number of faulty symbols per signal (N -dimension) is

$$E(L) = \sum_{i=0}^{M-1} P(\underline{s}_i) E(L | \underline{s}_i) = \frac{1}{M} \sum_i E(L | \underline{s}_i). \quad (50)$$

We define then the following sets, $E_{i,l} := \{\cup \underline{s}_j \mid d_H(\underline{s}_i, \underline{s}_j) = l\}$, $0 \leq i \leq M-1$, $0 \leq l \leq N/2$, where $d_H(\underline{s}_i, \underline{s}_j)$ is the Hamming distance between the signals \underline{s}_i and \underline{s}_j in terms of its constituent 2D symbols. We can, therefore, write

$$E(L | \underline{s}_i) = \sum_{l=0}^{N/2} l P(\text{dec } \underline{s} \in E_{i,l}). \quad (51)$$

Defining the associated index sets as $T_{i,l} := \{j \in 0, 1, \dots, M-1 \mid d_H(\underline{s}_i, \underline{s}_j) = l\}$, we thus have:

$$P(\text{dec } \underline{s} \in E_{i,l}) = P\left(\bigcup_{j \in T_{i,l}} \text{dec } \underline{s}_j\right). \quad (52)$$

Again, we can build a partition on the index set according to:

$$T_{i,l}^d = \{j \in 0, 1, \dots, M-1 \mid d_H(\underline{s}_i, \underline{s}_j) = l \cap d^2(\underline{s}_i, \underline{s}_j) = d^2\}, \quad d^2 > 0. \quad (53)$$

Applying this partition on the index set yields $E_{i,l} = \bigcup_{d=d_{\min}} \bigcup_{j \in T_{i,l}^d} \underline{s}_j$. Now, the union bound results in

$$E(L | \underline{s}_i) = \sum_{l=0}^{N/2} \sum_{d=d_{\min}}^{d_{\max}} \sum_{j \in T_{i,l}^d} l P(\underline{s}_i \rightarrow \underline{s}_j). \quad (54)$$

We shall denote by $N_i(d^2, l) := |T_{i,l}^d|$ the cardinality of the set $T_{i,l}^d$. Certainly, $\sum_{l=0}^{N/2} N_i(d^2, l) = N_i(d^2)$.

By the same token as Theorem 1, it is found that

$$E(L | \underline{s}_i) \leq \sum_{l=0}^{N/2} \sum_{d=d_{\min}}^{d_{\max}} l N_i(d^2, l) e(d^2) \quad (55)$$

and finally,

$$E(L) \leq \sum_{l=0}^{N/2} \sum_{d=d_{\min}}^{d_{\max}} l N_{av}(d^2, l) e(d^2) \quad (56)$$

where $N_{av}(d^2, l) := \frac{1}{M} \sum_{i=0}^{M-1} N_i(d^2, l)$ is the average number of pairs of signal having both squared Euclidean distance d^2 and l distinct symbols.

By using the exponential bound on the $\text{erfc}(\cdot)$ function, (56) can be rewritten as

$$E(L) \leq 1/2 \sum_{d=d_{\min}}^{d_{\max}} \sum_{l=1}^{N/2} l N_{av}(d^2, l) D^{d^2/d_0^2} \Big|_{D=D_0}. \quad (57)$$

On the other hand, it can be seen from definition 5 that

$$\frac{\partial T(D, L; \Omega)}{\partial L} = \sum_{d=d_{\min}}^{d_{\max}} \sum_{l=1}^{N/2} l N_{av}(d^2, l) D^{d^2/d_0^2} L^{l-1}. \quad (58)$$

Comparing (58) and the right hand side of (57), we recognize that

$$E(L) \leq 1/2 \left. \frac{\partial T(D, L; \Omega)}{\partial L} \right|_{L=1, D=D_0}. \quad (59)$$

Furthermore, the symbol error probability can be found to be $P_e(2) = \frac{E(L)}{N/2}$ so the claimed bound follows. **Q.E.D.**

LEMMA 13. The upper bound on the 2-dimensional error probability (Theorem 12) is less than or equal to that on the N -dimensional error probability (Lemma 7), that is,

$$\frac{1}{N} \left. \frac{\partial T(D, L; \Omega)}{\partial L} \right|_{L=1, D=D_0} \leq 1/2 \left. T(D, L; \Omega) \right|_{L=1, D=D_0}.$$

proof. Using definition 5, bound (35) on the N -dimensional error rate can be rewritten as

$$P_e(N) \leq 1/2 \left. T(D, L; \Omega) \right|_{L=1, D=D_0}. \quad (60)$$

Furthermore, it can be noted from (58) that

$$\left. \frac{\partial T(D, L; \Omega)}{\partial L} \right|_{L=1} \leq N/2 \sum_{d=d_{\min}} \sum_{l=1} N_{av}(d^2, l) D^{d^2/d_0^2} \quad (61)$$

since $l \leq N/2$ always holds. Accordingly, the asserted inequality follows by dividing both members by N . **Q.E.D.**

In the following, these results are applied to multidimensional constellations with random coded modulation.

THEOREM 14. The symbol error rate of multidimensional random coded constellations is upper bounded by

$$P_e(2) \leq 1/2 T_2(D_0)^{N/2} \cdot \frac{[T_2(D_0) - 1]}{T_2(D_0)}.$$

proof. First, we consider the constituent 2-dimensional constellation Ω_2 . We introduce then the two-variable transfer function in agreement with

$$T(D, L; \Omega_2) = L T(D; \Omega_2). \quad (62)$$

Consequently, the associated complete transfer function is

$$T_2(D, L) = 1 + L [T_2(D) - 1]. \quad (63)$$

Obviously, one has $T_2(D, 1) = T_2(D)$ as expected. We go further to N -dimensions and calculate the complete transfer function

$$T_N(D, L) = [T_2(D, L)]^{N/2} \quad (64)$$

so we obtain an expression of the transfer function in terms of its 2D constituent:

$$T(D, L; \Omega) = [T_2(D, L)]^{N/2} - 1. \quad (65)$$

In order to apply Theorem 12, we take the partial derivative of the above function

with respect to L , resulting in

$$\frac{\partial T(D, L; \Omega)}{\partial L} = \frac{N}{2} [T_2(D, L)]^{N/2-1} \frac{\partial T_2(D, L)}{\partial L}. \quad (66)$$

Furthermore, it can be seen from (63) that

$$\left. \frac{\partial T_2(D, L)}{\partial L} \right|_{L=1} = T_2(D) - 1. \quad (67)$$

Therefore, the result follows by inserting (67) into (66) and then applying Theorem 12. Q.E.D.

This can be illustrated by the example of the random coded constellation with a 4-point constituent 2D constellation. In that event, we have $T_2(D, L) = 1 + 2DL + 1D^2L$. For $N = 4$, for instance, it follows that $T_4(D, L) = 1 + 4DL + (2L + 4L^2)D^2 + 4D^3L^2 + 1D^4L^2$. The bound on the symbol error rate resulting from the previous theorem is $P_e(2) \leq D_0 + 2.5D_0^2 + 2D_0^3 + 0.5D_0^4$.

IV.2.5. ASYMPTOTIC BEHAVIOUR OF VERY LONG CODES

*"Virtually all large coded constellations are good
because they are essentially identical", H.M. de OLIVEIRA.*

Another question which should be addressed is about the famous statement "(almost) all codes are good", now thinking of the coded modulation context. We present below a straightforward theorem which clarifies this point.

THEOREM 15. (ALL CODED CONSTELLATIONS ARE GOOD): Given an arbitrary $\varepsilon > 0$, if we pick at random a particular coded (redundant) signal set #1 with length $N \rightarrow \infty$, then almost surely (a.s.) it is a good one (that is, $P_e^{\#1}(N) < \varepsilon$), provided $R < \bar{R}_0$.

proof. The theorem on the partial cut-off rate asserts that $\bar{P}_e(N) \leq \exp_2(-N[\bar{R}_0(N) - R])$. Applying then Definition 2, it follows readily that $\lim_{N \rightarrow \infty} \bar{P}_e(N) = 0$, provided $R < \bar{R}_0$. Therefore, recalling (14), we can write

$$\lim_{N \rightarrow \infty} \bar{P}_e(N) = \lim_{N \rightarrow \infty} \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L P_e^{\#l}(N) = 0. \quad (68)$$

Changing the order of the limits, we get

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \left[\lim_{N \rightarrow \infty} P_e^{\#l}(N) \right] = 0, \quad (69)$$

still assuming that $R < \bar{R}_0$.

On the other hand, given $\varepsilon > 0$, we define bad and good sets, respectively, by

$$B_\varepsilon(L) := \{ l \in 1, 2, \dots, L \mid \lim_{N \rightarrow \infty} P_\varepsilon^{\#l}(N) > \varepsilon \}$$

and

$$G_\varepsilon(L) := \{ l \in 1, 2, \dots, L \mid \lim_{N \rightarrow \infty} P_\varepsilon^{\#l}(N) \leq \varepsilon \}.$$

Thus, the bad set is an index set pointing out all the codes for which the ND-error probability is not arbitrarily small, as the dimensionality increases without limit. Of course, $(\forall \varepsilon) |B_\varepsilon(L)| + |G_\varepsilon(L)| = L$ since every code is either good or bad. Let B_ε (resp. G_ε) be the bad (resp. good) set when L increases indefinitely. Accordingly,

$$0 = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \left[\lim_{N \rightarrow \infty} P_\varepsilon^{\#l}(N) \right] \geq \lim_{L \rightarrow \infty} \frac{1}{L} \{ |B_\varepsilon(L)| \cdot \varepsilon + |G_\varepsilon(L)| \cdot 0 \}. \quad (70)$$

Furthermore, $|B_\varepsilon(L)|/L$ converges a.s. (by the strong law of large numbers) to the probability of taking a bad code, $P(B_\varepsilon)$, that is, $\lim_{L \rightarrow \infty} |B_\varepsilon(L)| \cdot L^{-1} = P(B_\varepsilon)$ a.s..

Hence, we find out $P(B_\varepsilon) \cdot \varepsilon \leq 0$ a.s. which implies $P(B_\varepsilon) = 0$ a.s. Thus, except for a set of measure zero, all the codes have vanishing error probability as N goes to infinity. Q.E.D.

Indeed, like the random codes, random coded constellations are not uniformly good but similar remarks ([15], p. 39) can be made.

COROLLARY 16. Almost all coded constellations exhibit virtually the same performance, provided their number of dimensions is very large.

proof. Consider two arbitrary codes, denoted by the superscripts $\#i$ and $\#j$, respectively. We shall prove that almost surely $|P_\varepsilon^{\#i}(N) - P_\varepsilon^{\#j}(N)| \leq \varepsilon$ when $N \rightarrow \infty$. Conversely, assume that $|P_\varepsilon^{\#i}(N) - P_\varepsilon^{\#j}(N)| > \varepsilon$. Without loss of generality it may be assumed that $P_\varepsilon^{\#i}(N) > P_\varepsilon^{\#j}(N)$ so that $P_\varepsilon^{\#i}(N) > \varepsilon + P_\varepsilon^{\#j}(N) > \varepsilon$. Hence, at least one of the two codes must belong to the bad set when n increases indefinitely. Therefore, $P(|P_\varepsilon^{\#i}(N) - P_\varepsilon^{\#j}(N)| > \varepsilon) \leq P(i \cup j \in B_\varepsilon) \leq 2P(B_\varepsilon)$ which goes a.s. to zero and our claim follows. Q.E.D.

Any open interval on the real line is denoted henceforward by an open ball of center K and radius κ , $B(K, \kappa) = \{\delta^2 \mid |\delta^2 - K| < \kappa\}$, and its complementary event by $B^c(K, \kappa)$.

LEMMA 17. Almost all constellations exhibit the distance hardening phenomenon as their dimensionality increases indefinitely.

proof. Let $K = E(\delta^2)$ and an arbitrary $\kappa > 0$. We want to show that a very long code picked at random satisfies $P^{\#i}(|\delta^2 - E(\delta^2)| \geq \kappa) = P^{\#i}(\delta^2 \in B^c(K, \kappa)) \rightarrow 0$ a.s.. From (16) and Definition 1, we have

$$\bar{P}(\delta^2) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L P^{\#l}(\delta^2). \quad (71)$$

Summing up for any $\delta^2 \in B^c(K, \kappa)$ and changing the order of summations in the right-hand side, we get

$$\bar{P}(\delta^2 \text{ in } B^c(K, \kappa)) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L P^{\#l}(\delta^2 \text{ in } B^c(K, \kappa)) . \quad (72)$$

When N increases without limit, Theorem 11 implies that $\lim_{N \rightarrow \infty} \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L P^{\#l}(\delta^2 \text{ in } B^c(K, \kappa)) = 0$.

Now the proof follows quite close to that of Theorem 15. We define here the bad and good sets to be, respectively,

$$B_L(\epsilon) := \{l=1, 2, \dots, L \mid \lim_{N \rightarrow \infty} P^{\#l}(\delta^2 \text{ in } B^c(K, \kappa)) > \epsilon\}$$

and

$$G_L(\epsilon) := \{l=1, 2, \dots, L \mid \lim_{N \rightarrow \infty} P^{\#l}(\delta^2 \text{ in } B^c(K, \kappa)) \leq \epsilon\} .$$

The proof is then completed by the same reasoning as for Theorem 15. **Q.E.D.**

DEFINITION 6. Given $\epsilon > 0$, two discrete distributions $P^{\#i}(x)$ and $P^{\#j}(x)$ are said to be ϵ -indistinguishable iff one has

$$(\forall K \in R) (\forall \kappa > 0) \quad |P^{\#i}(x \text{ in } B(K, \kappa)) - P^{\#j}(x \text{ in } B(K, \kappa))| < \epsilon .$$

We can see the ϵ -indistinguishability as a binary operation between distributions which will be denoted by \sim_ϵ . Further properties of this relation are presented in Appendix B.

DEFINITION 7. (QUASI-IDENTICAL CODES): Two coded signal sets denoted by superscripts $\#i$ and $\#j$ are said to be identical within ϵ (quasi-identical) iff their normalized squared distance distribution are ϵ -indistinguishable.

The relation \sim_ϵ asymptotically induces an equivalence relation among codes as N increases (see appendix B). In the following we show that all large constellations become equivalent in the above sense. We adopt Sup (resp. Inf) to denote the least upper bound (resp. greatest lower bound).

THEOREM 18. For any $\epsilon > 0$, almost all codes (coded constellations) are quasi-identical to the average code (random code), provided their dimensionality N is very large.

proof. We pick at random a redundant constellation (superscript $\#i$) on the assumption that the dimension is high enough. We intend to apply P2 (Appendix B) to show that $P^{\#i}(\delta^2) \sim_\epsilon \bar{P}(\delta^2)$. Given any $\kappa > 0$, we begin by splitting R^+ into two sets, namely

$$R_O^+ := \{K \in R^+ \mid E(\delta^2) \in B^c(K, \kappa)\} \quad \text{and} \quad R_I^+ := \{K \in R^+ \mid E(\delta^2) \in B(K, \kappa)\} .$$

Now we must evaluate

$$\max_{S=\{R_0^*, R_i^*\}} \sup_{K \in S} |P^{\#i}(\delta^2 \text{ in } B(K, \kappa)) - \bar{P}(\delta^2 \text{ in } B(K, \kappa))| = \max_{S=\{R_0^*, R_i^*\}} \sup_{K \in S} F_{\kappa}(K) \quad (73)$$

Firstly, we carry out the supremum of $F_{\kappa}(K)$ on R_0^+ . In this case we have

$$\sup_{K \in R_0^+} F_{\kappa}(K) \leq \sup_{K \in R_0^+} P^{\#i}(\delta^2 \text{ in } B(K, \kappa)) + \sup_{K \in R_0^+} \bar{P}(\delta^2 \text{ in } B(K, \kappa)) . \quad (74)$$

It follows from the squared distance hardening properties (Theorem 15 and Lemma 17) that

$$\sup_{K \in R_0^+} F_{\kappa}(K) \leq \varepsilon/2 + \varepsilon/2 \leq \varepsilon \quad \text{a.s.} \quad (75)$$

Secondly, we carry out the supremum of $F_{\kappa}(K)$ on R_i^+ . Now we use the fact that

$$\sup_{K \in R_i^+} F_{\kappa}(K) \leq \sup_{K \in R_i^+} |1 - P^{\#i}(\delta^2 \text{ in } B(K, \kappa))| + \sup_{K \in R_i^+} |1 - \bar{P}(\delta^2 \text{ in } B(K, \kappa))| \quad (76)$$

yielding

$$\sup_{K \in R_i^+} F_{\kappa}(K) \leq (1 - \inf_{K \in R_i^+} P^{\#i}(\delta^2 \text{ in } B(K, \kappa))) + (1 - \inf_{K \in R_i^+} \bar{P}(\delta^2 \text{ in } B(K, \kappa))) \quad (77)$$

Once again, hardening properties lead to

$$\sup_{K \in R_i^+} F_{\kappa}(K) \leq \varepsilon/2 + \varepsilon/2 \leq \varepsilon \quad \text{a.s.} \quad (78)$$

The proof is completed by combining (75) and (78) with (73). **Q.E.D.**

This theorem has the following corollary:

COROLLARY 19. Almost all codes become quasi-identical for any $\varepsilon > 0$ as N increases without limit.

proof. Consequence of both the above theorem and the asymptotic transitivity property on P3 (appendix B). **Q.E.D.**

IV.2.6. A GALLAGER TYPE BOUND FOR CODED MODULATION

In this section we intend to derive upper bounds on the error probability of finite-dimensional coded constellations with equally likely signals which are, in many ways, similar to those by Gallager [16]. They therefore generalize bounds established in the previous sections. In particular, attainable limits on the performance of coded modulation signals are studied. A family of generalized union bound is introduced which results in an E_0 function that allows to support rates beyond the cutoff rate. We define a capacity (and a cutoff rate) of coded finite-dimensional signal sets as the rate that cannot be exceeded by these systems provided that the error probability can be made arbitrarily small.

The idea here is to generalize the union bound and then apply these results to evaluate performance of combined coding and modulation schemes previously described. First, we shall need some useful inequalities given by

PROPOSITION I (INEQUALITIES). Let I be a countable index set. Any sequence $\{x_i\}_{i \in I}$ of real numbers verifies the following inequalities

$$\begin{aligned} (i) \quad \sum_{i \in I} |x_i| &\leq \left[\sum_{i \in I} |x_i|^{\frac{1}{1+p}} \right]^{p+1} \quad \forall p \geq 0, \\ (ii) \quad \sum_{i \in I} |x_i| &\geq \left[\sum_{i \in I} |x_i|^{\frac{1}{1+p}} \right]^{p+1} \quad \forall p: -1 < p \leq 0. \quad \blacksquare \end{aligned}$$

These are well known inequalities (see e.g. [18], p.197). Nevertheless, we propose an alternative proof of them in Appendix C.

PROPOSITION II. Let $\{x_i\}_{i \in I}$ be a sequence of real numbers and let $\{n_i\}_{i \in I}$ denote a sequence of positive integers, $n_i \geq 0$. Then for any $p > 0$, we have the inequality

$$\sum_{i \in I} n_i |x_i| \leq \left[\sum_{i \in I} n_i |x_i|^{\frac{1}{p+1}} \right]^{p+1}$$

Proposition I above may be applied in conjunction with the union bound in order to derive a family of upper bounds on the probability of a union of events. To begin with, let $\{P(A_j)\}_{j \in I}$ be the respective probabilities of events $\{A_j\}_{j \in I}$.

Consequently, we can easily infer from part (i) of Proposition I that, for any $p \geq 0$,

$$\sum_{j \in I} P(A_j) \leq \left[\sum_{j \in I} P(A_j)^{\frac{1}{1+p}} \right]^{p+1} \quad (79)$$

So we have established the following family of bounds:

PROPOSITION III (GENERALIZED UNION BOUND). Let $\{A_j\}_{j \in I}$ be an ensemble of (not necessarily disjoint) events. Then the following bounds hold for any $p \geq 0$:

$$P\left(\bigcup_{j \in I} A_j\right) \leq \left[\sum_{j \in I} P(A_j)^{\frac{1}{1+p}} \right]^{p+1} \quad \blacksquare$$

Clearly, the union bound is the particular case corresponding to $p = 0$. Also, we get some kind of Bhattacharyya-union bound for $p = 1$, namely

$$P\left(\bigcup_{j \in I} A_j\right) \leq \sum_{i \in I} \sum_{j \in I} \sqrt{P(A_i)P(A_j)}. \quad (80)$$

Furthermore, we mention that most bounds in this family are strictly weaker than the

union bound. However, this is not inconvenient for one who wishes to study the performance as rate increases towards capacity. We discuss now this point.

There are two different ways to think about powerful bounds like Gallager's.

First, we may say that the Bhattacharyya-union bound does not allow to achieve capacity because it is not a very good upper bound (it is too loose), so we need tighter upper bounds. This seems to be the thinking of most researchers. They consider Gallager's bound tighter than the Bhattacharyya-union bound although both coincides in the range of rates where the latter is remains valid. The former also holds in a considerable range beyond the cutoff rate. However, considering the family of bounds as a function of ρ without carrying out optimization with respect to this parameters, most bounds are weaker.

As a second approach, we examine the close relationship between the transmission rate R and the error performance $P_e(N)$. Indeed, rising the transmission rate increases the error probability. Therefore, making severe constraints about the way in which the probability of error goes to zero, forbids us to handle very high rates. In other words, one should slack constraints about the way that the error probability goes to zero in order to allow higher rates. Hence, we need weaker bounds than the Bhattacharyya-union bound to achieve capacity. For very-high rates (above the critical rate), the Bhattacharyya-union bound is no longer valid, that is, there exist no codes having such a behaviour of decreasing error rate. In contrast, Gallager bounds show the existence of codes with vanishing error probability. Accordingly, we prefer to use the term "more general bounds" rather than "tighter bounds", although both interpretations are correct.

Now we consider (coded) finite-dimensional constellations with M equally likely signals on an additive white Gaussian noise (AWGN) channel. Again, by abuse of language, a coded constellation will be referred to indistinctly as a code.

LEMMA 20. The N -dimensional (block) error probability $P_e(N)$ for a particular code, say with superscript $\#l$, is upper bounded by

$$P_e^{\#l}(N) \leq \exp_2 -N E_r^{\#l}(R, N) ,$$

where $E_r^{\#l}(R, N) := \sup_{\rho > 0} [-\rho R + E_r^{\#l}(\rho, N)]$ is a reliability function of the code, namely

$$E_0^{\#l}(\rho, N) := -\frac{1}{N} \log_2 \sum_{i=0}^{M-1} \left[\sum_{d=d_{\min}} P_i^{\#l}(d^2) e(d^2)^{\frac{1}{1+\rho}} \right]^{\rho+1} ,$$

where $e(d^2)$ denotes the pairwise error probability of any pair of signal points whose squared Euclidean distance is d^2 .

Some comments would be worthwhile. First, note that *each* code has an upper bound expressed in the above form. We are especially interested in examining bounds resulting from random coding which represents the average behaviour of the family of codes considered. Additionally, these bounds are based on the spectrum of the random code which is simpler to obtain. Secondly, we mention as a particular case a Bhattacharyya-union type bound which corresponds to let $\rho = 1$

$$P_e^{\#l}(N) \leq \exp_2 -N [R_0^{\#l}(N) - R],$$

where we have defined a code cutoff rate according to

$$R_0^{\#l}(N) := E_0^{\#l}(1, N).$$

It should be pointed out that this definition is close to the one of the conventional cut-off rate R_0 .

We feel that this bound is tight for very high dimensional codes, since it is an exponential bound on the (block) error probability provided that R does not exceed C . Such bounds are, therefore, interesting in order to analyse the asymptotic behaviour of very long codes. Indeed, they are not very tight for short codes. Moreover, these are bounds on the error probability of a *given* code $\#l$. However, it is quite difficult to calculate bounds for particular codes. Hence, we should consider the random coding which deals with all codes.

IV.2.7. CONCLUDING REMARKS

We applied Shannon's random coding argument to coded modulation. Upper bounds on the block and symbol error probability of finite-dimensional constellations have been developed. We defined a partial cut-off rate in terms of which we can upper bound the average error probability. An open problem is the relationship between the cut-off rate \bar{R}_0 and the conventional cut-off rate R_0 [16]. A more general upper bound than union bound should be used to allow for higher rates. Added to that, we mention that Hughes' cone bound [17] could be used instead of the union bound in order to improve the error probability estimation, especially for short codes.

Concerning the squared Euclidean distance distribution, parameters like its first two moments have been calculated. Finally, the asymptotic behaviour of such systems has been investigated. We have found that a hardening phenomenon occurs with the normalized squared Euclidean distance. We have also shown that virtually all large coded signal sets are good with probability asymptotically approaching one provided the rate does not exceed a critical value. The investigations carried out in this paper provide a better insight into the performance of multidimensional coded constellations.

APPENDIX A

DETERMINATION OF THE TRANSFER FUNCTION FOR THE CONSTITUENT 2D CONSTELLATION

We intend here to clarify the process for determining the finite theta series/transfer function of a constituent 2D constellation. The set of the distances between two arbitrary points in the Ω_2 constellation is denoted by $\Delta(\Omega_2)$. [A] Once the number of signal points with a particular squared distance $d^2 \in \Delta(\Omega_2)$ to a given point may vary with this point, we are led to consider a matrix $[N](d^2)$ which exhibits the number of points at a squared distance d^2 from *each* signal point. The average number of signal pairs with squared distance d^2 , $N_{av}(d^2)$, is found by averaging all the elements in the $[N](d^2)$ matrix. Thus, finite theta series and transfer functions can easily be determined according to Definitions 2 and 4. This process is better understood by considering representative cases: the following concerns particular signal sets, namely square/cross M -ary QAM. Moreover, we assume $d_0 = 2$ for convenience.

CASE 1: $|\Omega_2| = 4$. The complex constellation (4-ary QAM) is represented by the following matrix of coordinates

$$[\Omega_2] = \begin{bmatrix} (-1, 1) & (1, 1) \\ (-1, -1) & (1, -1) \end{bmatrix}.$$

We have $\Delta(\Omega_2) = \{0, d_0^2, 2d_0^2\}$, and the neighbourhood matrices are

$$[N](0) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad [N](d_0^2) = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \quad [N](2d_0^2) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

The finite theta series for this signal set is, therefore, expressed by $\Theta_{\Omega_2}(z) = 1 + 2w^{d_0^2} + 1w^{2d_0^2}$ so the complete transfer function is $T_2(D) = 1 + 2D + 1D^2$. Additionally, we notice that $[N](0)$ has all its entries equal to unity for any kind of signal set, so $N_{av}(0) = 1$ as expected.

CASE 2: $|\Omega_2| = 16$. The complex constellation matrix for 16-ary QAM is given by

$$[\Omega_2] = \begin{bmatrix} (-3, 3) & (-1, 3) & (1, 3) & (3, 3) \\ (-3, 1) & (-1, 1) & (1, 1) & (3, 1) \\ (-3, -1) & (-1, -1) & (1, -1) & (3, -1) \\ (-3, -3) & (-1, -3) & (1, -3) & (3, -3) \end{bmatrix}.$$

It can be readily checked that

$$\Delta(\Omega_2) = \{0, d_0^2, 2d_0^2, 4d_0^2, 5d_0^2, 8d_0^2, 9d_0^2, 10d_0^2, 13d_0^2, 18d_0^2\}.$$

The respective neighbourhood matrices for every $d^2 \in \Delta(\Omega_2)$ are:

$$[N](d_0^2) = \begin{bmatrix} 2 & 3 & 3 & 2 \\ 3 & 4 & 4 & 3 \\ 3 & 4 & 4 & 3 \\ 2 & 3 & 3 & 2 \end{bmatrix} \quad [N](2d_0^2) = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & 2 \\ 2 & 4 & 4 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix} \quad [N](4d_0^2) = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix}$$

$$\begin{aligned}
[N](5d_0^2) &= \begin{bmatrix} 2 & 3 & 3 & 2 \\ 3 & 4 & 4 & 3 \\ 3 & 4 & 4 & 3 \\ 2 & 3 & 3 & 2 \end{bmatrix} & [N](8d_0^2) &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} & [N](9d_0^2) &= \begin{bmatrix} 2 & 1 & 1 & 2 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix} \\
[N](10d_0^2) &= \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix} & [N](13d_0^2) &= \begin{bmatrix} 2 & 1 & 1 & 2 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix} & [N](18d_0^2) &= \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.
\end{aligned}$$

The values of $N_{av}(d^2)$ for $d^2 \in \Delta(\Omega_2)$ are computed by averaging all the elements in the corresponding matrix, as previously described.

A SECOND EXAMPLE:

$$[\Omega_2] = \begin{bmatrix} (-, -) & (-3, 5) & (-1, 5) & (1, 5) & (3, 5) & (-, -) \\ (-5, 3) & (-3, 3) & (-1, 3) & (1, 3) & (3, 3) & (5, 3) \\ (-5, 1) & (-3, 1) & (-1, 1) & (1, 1) & (3, 1) & (5, 1) \\ (-5, -1) & (-3, -1) & (-1, -1) & (1, -1) & (3, -1) & (5, -1) \\ (-5, -3) & (-3, -3) & (-1, -3) & (1, -3) & (3, -3) & (5, -3) \\ (-, -) & (-3, -5) & (-1, -5) & (1, -5) & (3, -5) & (-, -) \end{bmatrix}$$

It can be readily checked that

$$\Delta(\Omega_2) = \{0, d_0^2, 2d_0^2, 4d_0^2, 5d_0^2, 8d_0^2, 9d_0^2, 10d_0^2, 13d_0^2, 16d_0^2, 18d_0^2, 20d_0^2, 25d_0^2, 26d_0^2, 29d_0^2, 32d_0^2, 34d_0^2\}.$$

The respective neighborhood matrices for every $d^2 \in \Delta(\Omega_2)$ are:

$$\begin{aligned}
[N](d_0^2) &= \begin{bmatrix} - & 2 & 3 & 3 & 2 & - \\ 2 & 4 & 4 & 4 & 4 & 2 \\ 3 & 4 & 4 & 4 & 4 & 3 \\ 3 & 4 & 4 & 4 & 4 & 3 \\ 2 & 4 & 4 & 4 & 4 & 2 \\ - & 2 & 3 & 3 & 2 & - \end{bmatrix} & [N](2d_0^2) &= \begin{bmatrix} - & 2 & 2 & 2 & 2 & - \\ 2 & 3 & 4 & 4 & 3 & 2 \\ 2 & 4 & 4 & 4 & 4 & 2 \\ 2 & 4 & 4 & 4 & 4 & 2 \\ 2 & 3 & 4 & 4 & 3 & 2 \\ - & 2 & 2 & 2 & 2 & - \end{bmatrix} & [N](4d_0^2) &= \begin{bmatrix} - & 2 & 2 & 2 & 2 & - \\ 2 & 2 & 3 & 3 & 2 & 2 \\ 2 & 3 & 4 & 4 & 3 & 2 \\ 2 & 3 & 4 & 4 & 3 & 2 \\ 2 & 2 & 3 & 3 & 2 & 2 \\ - & 2 & 2 & 2 & 2 & - \end{bmatrix} \\
[N](5d_0^2) &= \begin{bmatrix} - & 3 & 4 & 4 & 3 & - \\ 3 & 4 & 5 & 5 & 4 & 3 \\ 4 & 5 & 8 & 8 & 5 & 4 \\ 4 & 5 & 8 & 8 & 5 & 4 \\ 3 & 4 & 5 & 5 & 4 & 3 \\ - & 3 & 4 & 4 & 3 & - \end{bmatrix} & [N](8d_0^2) &= \begin{bmatrix} - & 1 & 2 & 2 & 1 & - \\ 1 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 3 & 3 & 2 & 2 \\ 2 & 2 & 3 & 3 & 2 & 2 \\ 1 & 1 & 2 & 2 & 1 & 1 \\ - & 1 & 2 & 2 & 1 & - \end{bmatrix} & [N](9d_0^2) &= \begin{bmatrix} - & 2 & 1 & 1 & 2 & - \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 2 & 2 & 2 & 2 & 1 \\ 1 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ - & 2 & 1 & 1 & 2 & - \end{bmatrix}
\end{aligned}$$

$$]_{(10d_0^2)} = \begin{bmatrix} 3 & 3 & 3 & 3 \\ 3 & 4 & 3 & 3 & 4 & 3 \\ 3 & 3 & 4 & 4 & 3 & 3 \\ 3 & 3 & 4 & 4 & 3 & 3 \\ 3 & 4 & 3 & 3 & 4 & 3 \\ 3 & 3 & 3 & 3 & 3 & 3 \end{bmatrix} \quad [N]_{(13d_0^2)} = \begin{bmatrix} 2 & 3 & 3 & 2 \\ 2 & 2 & 3 & 3 & 2 & 2 \\ 3 & 3 & 2 & 2 & 3 & 3 \\ 3 & 3 & 2 & 2 & 3 & 3 \\ 2 & 2 & 3 & 3 & 2 & 2 \\ 2 & 3 & 3 & 2 & 3 & 2 \end{bmatrix} \quad [N]_{(16d_0^2)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$]_{(17d_0^2)} = \begin{bmatrix} 3 & 2 & 2 & 3 \\ 3 & 2 & 2 & 2 & 2 & 3 \\ 2 & 2 & 0 & 0 & 2 & 2 \\ 2 & 2 & 0 & 0 & 2 & 2 \\ 3 & 2 & 2 & 2 & 2 & 3 \\ 3 & 2 & 2 & 3 & 3 & 3 \end{bmatrix} \quad [N]_{(18d_0^2)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad [N]_{(20d_0^2)} = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

$$]_{(25d_0^2)} = \begin{bmatrix} 3 & 2 & 2 & 3 \\ 3 & 2 & 0 & 0 & 2 & 3 \\ 2 & 0 & 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 & 0 & 2 \\ 3 & 2 & 0 & 0 & 2 & 3 \\ 3 & 2 & 2 & 3 & 3 & 3 \end{bmatrix} \quad [N]_{(26d_0^2)} = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 1 & 1 \end{bmatrix} \quad [N]_{(29d_0^2)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$[N]_{(32d_0^2)} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad [N]_{(34d_0^2)} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

APPENDIX B

PROPERTIES OF THE RELATION \sim_ϵ

We present here a few properties of the relation \sim_ϵ involving the definition of

ϵ -indistinguishable distributions.

P1. If $P^{#i}(x) \sim_{\epsilon_1} P^{#j}(x)$ then it follows that $P^{#i}(x) \sim_{\epsilon_2} P^{#j}(x)$ provided $\epsilon_2 > \epsilon_1$.

P2. A necessary and sufficient condition for two distributions to be ϵ -indistinguishable is that $\forall \kappa > 0$

$$\sup_{K \in \mathcal{R}} |P^{#i}(x \text{ in } B(K, \kappa)) - P^{#j}(x \text{ in } B(K, \kappa))| < \epsilon.$$

P3. The relation induced by ϵ -indistinguishable normalized squared distance distributions is asymptotically an equivalence relation, that is,

a) \sim_{ϵ} is reflexive, b) \sim_{ϵ} is symmetric, c) \sim_{ϵ} is asymptotically transitive.

proof. a) and b) are trivial. To prove c), given $\epsilon > 0$, there exists N large enough such that $P^{#i} \sim_{\epsilon/2} P^{#j}$ and $P^{#j} \sim_{\epsilon/2} P^{#k}$. Now $P^{#i} \sim_{\epsilon} P^{#j}$ and $P^{#j} \sim_{\epsilon} P^{#k}$ follows from P1. The asymptotic transitivity follows from P2 and the triangular inequality.

APPENDIX C

Proof of Proposition I. The reasoning used here follows Gallager [16]. Clearly, for any $\rho \geq 0$,

$$\Phi_i := \left[\frac{\sum_{j \in I} |x_j|^{1/(1+\rho)}}{|x_i|^{1/(1+\rho)}} \right]^{\rho} \geq 1, \quad (C1)$$

and therefore,

$$\sum_{i \in I} |x_i| \leq \sum_{i \in I} |x_i| \Phi_i = \sum_{i \in I} |x_i|^{\frac{1}{1+\rho}} \left(\sum_{j \in I} |x_j|^{\frac{1}{1+\rho}} \right)^{\rho} \quad (C2)$$

so part (i) follows.

Furthermore, we have from the above that, for any $0 < \beta \leq 1$,

$$\sum_{i \in I} |x_i| \leq \left[\sum_{i \in I} |x_i|^{\beta} \right]^{1/\beta}. \quad (C3)$$

Considering a given sequence $\{y_i\}_{i \in I}$, we apply then (C3) to $|x_i| = |y_i|^{1/\beta}$, which results in

$$\sum_{i \in I} |y_i|^{1/\beta} \leq \left[\sum_{i \in I} |y_i| \right]^{1/\beta} \quad \text{for any } 0 < \beta \leq 1. \quad (C4)$$

The proof of part (ii) is completed by letting $\beta = 1+\rho$ ($-1 < \rho \leq 0$). **Q.E.D.**

Proof of Proposition II. Define a new index set I' with cardinality $|I'| = (\sum_{i \in I} n_i)$ which contains I but where the element i is repeated n_i times. Now

$$\sum_{i \in I} n_i |x_i| = \sum_{i \in I'} |x_i| \leq \left[\sum_{i \in I'} |x_i|^{\frac{1}{1+p}} \right]^{p+1}, \quad (C5)$$

where the inequality follows from Proposition I. The proof is completed by changing the index set from I' to I . **Q.E.D.**

Proof of lemma 20. The error probability of a code $\#l$ can be evaluated in terms of the conditional error probability given that a particular signal \underline{s}_j is transmitted, i.e.,

$$P_e^{\#l}(N) = \sum_{i=0}^{M-1} \frac{1}{M} P(e | \underline{s}_i). \quad (C6)$$

The conditional error probability can be bounded by a (generalized) union bound as follows. Let $\underline{s}_j \rightarrow \underline{s}_j$ denote the event decoding \underline{s}_j when \underline{s}_j is sent and \underline{s}_j is considered the only alternative. We suppose that Maximum Likelihood Decoding (MLD) is carried out by the decoder which is optimal in the sense that it minimizes the error rate. We also consider an index set T_i^d defined as $T_i^d := \{j \in 0, 1, \dots, M-1 \mid j \neq i, d^2(\underline{s}_i, \underline{s}_j) = d^2\}$. By a union bound argument, we find out

$$P(e | \underline{s}_i) \leq \sum_{d=d_{\min}} \sum_{j \in T_i^d} P^{\#l}(\underline{s}_i \rightarrow \underline{s}_j), \quad (C7)$$

where $P^{\#l}(\underline{s}_i \rightarrow \underline{s}_j)$ is the pairwise error probability which solely depends on d^2 but not on j . We denote by $e(d^2)$ the pairwise error probability of any pair of signal points whose squared Euclidean distance is d^2 .

Let $N_i^{\#l}(d^2) := |T_i^d|$ denote the cardinality of the set T_i^d , that is, the number of signals at a squared distance d^2 from \underline{s}_i . Applying then Proposition II to (C7), we get

$$P(e | \underline{s}_i) \leq \sum_{d=d_{\min}} N_i^{\#l}(d^2) e(d^2) \leq \left[\sum_{d=d_{\min}} N_i^{\#l}(d^2) e(d^2)^{\frac{1}{1+p}} \right]^{p+1} \quad (C8)$$

Thus, a bound on the block error probability of a code can be found by inserting (C8) into (C6). This can be rewritten as

$$P_e^{\#l}(N) \leq M^p \sum_{i=0}^{M-1} \left[\sum_{d=d_{\min}} P_i^{\#l}(d^2) e(d^2)^{\frac{1}{1+p}} \right]^{p+1}, \quad (C9)$$

where $P_i^{\#l}(d^2) := N_i^{\#l}(d^2)/M$ is the (conditional) distance spectrum of the code $\#l$, still assuming that $p \geq 0$.

We now define a so-called Gallager function $E_0^{\#l}(p, N)$, $0 \leq p < \infty$, according to

$$E_0^{\#l}(\rho, N) := -\frac{1}{N} \log_2 \sum_{i=0}^{M-1} \left[\sum_{d=d_{\min}} P_i^{\#l}(d^2) e^{(d^2) \frac{1}{1+\rho}} \right]^{\rho+1} \quad (C10)$$

(compare, for instance, with the $E_0(\cdot)$ function defined by Gallager on a discrete Memoryless Channel ([16])).

Remembering that $M = \exp_2(NR)$, it follows promptly that

$$P_e^{\#l}(N) \leq \exp_2 -N [-\rho R + E_0^{\#l}(\rho, N)], \quad 0 \leq \rho < \infty. \quad (C11)$$

Finally, we can choose the tightest bound by maximizing the term into brackets in the above expression according to

$$E_r^{\#l}(R, N) := \sup_{\rho \geq 0} [-\rho R + E_0^{\#l}(\rho, N)], \quad (C12)$$

and the proof follows. **Q.E.D.**

REFERENCES

- [1] C.E. Shannon, "Probability of Error for Optimal Codes in a Gaussian Channel", *BSTJ.*, **38**, n° 3, May 1959, pp. 611-656
- [2] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals", *IEEE Trans. Info. Theory*, **IT-28**, n° 1, 1982, pp. 55-67
- [3] K.M. Cheung, "Identities and Approximations for the Weight Distributions of Q-ary Codes", *IEEE Trans. Info. Theory*, **IT-36**, n° 5, Sept. 1990, pp. 1149-1153
- [4] G. Battail, H. Magalhães de Oliveira and Zhang Weidong, "Coding and Modulation for the Gaussian Channel, in the Absence or in the Presence of Fluctuations", *Proc. Int. Colloquium on Coding Theory and Applications*, EURO-CODE'90, Udine, Italy, Nov. 1990
- [5] A.J. Viterbi, "Convolutional Codes and their Performance in Communications Systems", *IEEE Trans. Commun. Technol.*, **COM-19**, n° 5, 1971, pp. 751-772
- [6] E. Zehavi and J.K. Wolf, "On the Performance Evaluation of Trellis Codes", *IEEE Trans. Info. Theory*, **IT-33**, n° 2, Mar. 1987, pp. 196-202
- [7] G. Battail, "Coding for the Gaussian Channel: the Promise of Weighted-Output Decoding", *Int. J. Satellite Commun.*, **7**, 1989, G. Battail, "Coding for the Gaussian Channel: the Promise of Weighted-Output Decoding", *Int. J. Satellite Commun.*, **7**, 1989, pp. 183-192
- [8] G. Battail, "Construction Explicite de Bons Codes Longs", *Ann. Télécomm.*, **44**, n° 7-8, July-Aug. 1989, pp. 392-404

- [9] H. Magalhães de Oliveira and G. Battail, "A Capacity Theorem for Lattice Codes on Gaussian Channels", *Proc. IEEE/SBT Int. Telecomm. Symposium, ITS'90*, Sept. 1990, Rio de Janeiro, Brazil, pp. 1.2.1-1.2.5
- [10] G.D. Forney and L-F. Wei, "Multidimensional Constellations - Part I: Introduction, Figures of Merit, and Generalized Cross Constellations", *IEEE J. Select. Areas Commun.*, SAC 7, n° 6, Aug. 1989, pp. 877-892
- [11] H. Magalhães de Oliveira and G. Battail, "On Generalized Constellations and the Opportunistic Secondary Channel", submitted.
- [12] N.J.A. Sloane, "Tables of Sphere Packings and Spherical Codes", *IEEE Trans. Info. Theory*, IT-27, n° 3, 1981, pp.327-338
- [13] J. H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, NY: Springer-Verlag, 1988
- [14] E.F. Beckenbach (Ed.), *Applied Combinatorial Mathematics*, NY: Wiley, 1964
- [15] P. Elias, "Coding for Noisy Channels", *IRE Conv. Rec.*, Mar. 1955, pp. 37-46
- [16] R.G. Gallager, *Information Theory and Reliable Communication*, NY: Wiley, 1968
- [17] B. Hughes, "On the Error Probability of Signals in Additive White Gaussian Noise", *IEEE Trans. Info. Theory*, IT-37, n° 1, 1991, pp.151-155
- [18] A.J. Viterbi and J.K. Omura, *Principles of Digital Communication and Coding*, NY: McGraw-Hill, 1979
- [19] J.N. Pierce, "Limit Distribution of the Minimum Distance of Random Linear Codes", *IEEE Trans. Info. Theory*, IT-13, n° 4, 1967, pp. 595-599

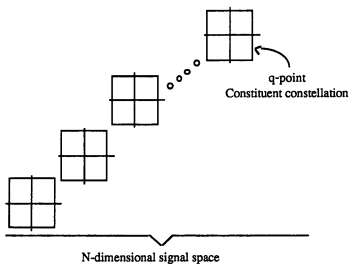
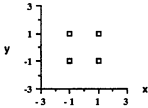
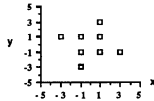


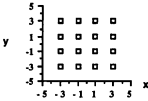
Figure 1. N -dimensional constellation created by taking the Cartesian product of a constituent 2-dimension constellation with itself $N/2$ times.

a) $M=4$

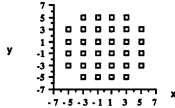
$$T_Z(D)=1+2D+1D^2$$

b) $M=8$

$$T_Z(D)=1+2.0D+1.5D^2+1.0D^4+2.0D^5+0.5D^{10}$$

c) $M=16$

$$T_Z(D)=1+3D+2.25D^2+2D^4+3D^5+1D^8+1D^9+1.5D^{10}+1D^{13}+0.25D^{16}$$

d) $M=32$

$$T_Z(D)=1+3.25D+2.875D^2+2.5D^4+4.5D^5+1.75D^8+1.75D^9+3.25D^{10}+2.5D^{13}+1D^{16}+2D^{17}+0.875D^{18}+1.5D^{20}+1.5D^{25}+0.75D^{28}+0.5D^{29}+0.25D^{32}+0.25D^{34}$$

Figure 2. M -point QAM constituent 2-dimensional constellations and their respective 2-dimensional transfer functions

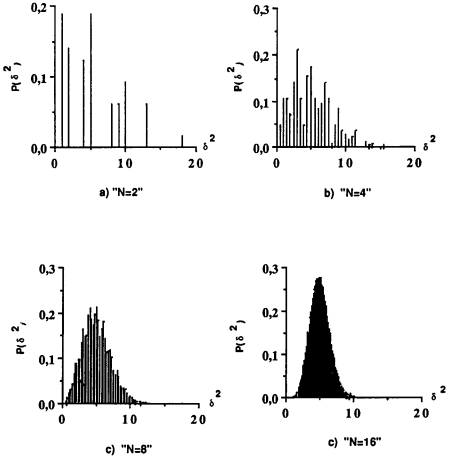


Figure 3. Evolution of (normalized) squared Euclidean distance spectrum of random coded constellations based on a 16-QAM constituent constellation for increasing dimensionality N .

**CONCLUSIONS
ET
PERSPECTIVES**

CONCLUSIONS.

L'évolution vertigineuse des débits exige que les systèmes numériques disposent de schémas de codage de canal de plus en plus performants. Les modulations codées sont apparues comme une réponse à ce phénomène et constituent maintenant un acquis définitif.

Les principaux résultats obtenus au cours de cette thèse peuvent être résumés ainsi:

- Un algorithme de détection pour la modulation MAQ au sens étroit.

- L'utilisation d'un canal secondaire opportuniste pour des constellations bidimensionnelles en croix généralisées ainsi que des bornes sur ses performances.

- L'étude du codage en treillis combiné à la modulation MDP différentielle.

- La démonstration sans équivoque de l'existence de moyens déterministes pour s'approcher de la capacité.

- La détermination d'une fonction de fiabilité pour les codes de réseau sur le canal gaussien.

- Des bornes supérieure et inférieure sur les performances de codes de réseau en fonction de leurs paramètres ainsi que leurs limites asymptotiques.

- Une discussion et une conjecture sur les rapports entre la capacité et les problèmes d'empilement de sphères et de recouvrement de l'espace euclidien (questions fondamentales pour éclairer la nature de la capacité).

- L'étude de la modulation codée aléatoire selon la méthode introduite par Shannon.

- Le calcul du spectre de distance euclidienne des constellations codées aléatoires à partir de la série θ finie.

- La mise en évidence du phénomène de durcissement de la distance euclidienne normalisée dont on montre qu'il est l'équivalent, pour la distance euclidienne, de la borne inférieure de Varshamov-Gilbert pour la distance de Hamming.

A l'avenir, il serait intéressant aussi d'étudier soigneusement le comportement des régions de Voronoi de réseaux dont le nombre de dimensions est très élevé. On remarque l'importance du fait que "*la presque totalité du volume d'une région convexe est concentrée sur sa surface*". Néanmoins, la connaissance des phénomènes dans un espace euclidien à grand nombre de dimensions est loin d'être une tâche aisée. Leurs propriétés sont très différents de celles de l'espace que nous habitons. Ces espaces sont, en effet, de véritables boîtes à surprises où beaucoup de phénomènes échappent à notre sens (voir, p.ex., le chapitre 9 de [HAM 1980]).

Une perspective digne d'attention serait l'interprétation des codes MCT comme des empilements de nombre de dimensions infini et invariants dans le temps. Il serait fort utile, d'un point de vue théorique, de démontrer un théorème de capacité pour le codage-modulation en treillis aboutissant à une expression *explicite* de leur fonction de fiabilité dans un canal à bruit gaussien additif, comme nous l'avons fait pour les réseaux. Bien qu'elle paraisse une constatation presque évidente, la supériorité des codes MCT sur les codes de réseau devrait exiger une analyse plus profonde, démontrant et quantifiant ainsi leur supériorité sur des constellations codées de dimension finie.

Rien n'empêche par ailleurs d'envisager d'appliquer le raisonnement utilisé pour obtenir le spectre de distance euclidienne des constellations codées (section IV.2) à la détermination de celui des modulations codées en treillis aléatoires.

D'autre part, l'emploi du choix aléatoire s'est montré une stratégie très riche et toujours utile pour l'étude des limites ultimes (codage de canal, minimisation de circuits booléens, etc). Il se pourrait que les idées exposées dans cette thèse puissent servir à trouver une borne sur la densité limite de réseaux à grand nombre de dimensions. A cet égard, il conviendrait de rappeler que l'empilement aléatoire à trois dimensions [BER 1969] présente une densité de 0,64 alors qu'elle vaut 0,74 pour les empilements les plus denses connus (p.ex. le réseau cubique à faces centrées, empilement hexagonal rapproché). On peut s'attendre à une diminution de l'écart entre la densité d'un empilement aléatoire et celle du meilleur réseau au fur et à mesure que le nombre des dimensions croît.

De même que l'on a généralisé la série θ aux constellations bornées, il est possible de définir leur densité comme la fraction de la région bornée recouverte par des sphères. La détermination d'un moyen de calculer la densité d'un réseau (constellation) en fonction de sa série θ (finie) pourrait aboutir à un calcul de la densité des constellations/empilements aléatoires. On pose alors la question de l'existence d'une limite à la grandeur ci-dessous et, si la réponse est positive, de son rapport avec la densité de centres :

$$\delta(\Lambda) = \delta(\Lambda:Z^n) \stackrel{?}{=} \frac{\Theta_{\Lambda}(0)}{\Theta_{Z^n}(0)}.$$

Si la modulation codée paraît d'une extrême fécondité, on constate tout même une certaine difficulté à progresser à mesure que l'on considère des codes très longs. Cette difficulté n'est pas encore vraiment ressentie à présent parce qu'elle apparaîtrait dans la conception de codes beaucoup trop longs pour que la technologie actuelle en permette la réalisation. D'ailleurs, un changement profond dans la conception et la philosophie d'une théorie fortement établie n'est obtenue qu'après une longue démarche. Dans notre contexte, rappelons les réticences de la communauté scientifique à accepter les idées de Shannon et d'Ungerboeck.

Il paraît exister une incompatibilité fondamentale entre les bons codes (au sens $d/n \rightarrow 0$) et leur structure mathématique [COF&GOO 1990]. Cela expliquerait pourquoi on ne trouve pas de bons codes. A ce propos, on pourrait proposer la comparaison suivante: faire correspondre les codes structurés aux nombres rationnels et les codes sans aucune structure à l'ensemble des irrationnels. Ayant associé la mesure de Lebesgue à l'intervalle $(0,1)$, la probabilité de choisir au hasard un nombre irrationnel (code non structuré) tend vers 1 alors que les nombres (codes) que nous concevons sont normalement rationnels (structurés). Il est donc facile de concevoir des codes structurés - probablement mauvais au sens de la distance minimale- quoique la presque totalité des codes soient bons. La démonstration de l'existence de bons codes de réseau n'est probablement pas en contradiction avec la discussion ci-dessus. En effet, si ces codes fortement structurés sont bons, c'est seulement en ce sens que leur probabilité d'erreur est petite. Cependant, ils ne sont probablement pas bons au sens $d/n \rightarrow 0$. La figure 7 (III.2) paraît traduire les principaux résultats de cette thèse. En interdisant quelques points, on pourrait obtenir des codes bons même au sens de la distance minimale (cette opération consiste à transformer un code bon en un code uniformément bon [ELI 1955]) en ne diminuant que très peu le débit d'information. Mais le prix à payer serait alors la perte de la riche structure mathématique. Battail propose en quelque sorte d'oublier le critère de distance minimale et de traiter des codes bons (*mais non uniformément bons*) pour conserver une structure favorable au décodage.

Ainsi, les contributions de ces travaux confirment l'intérêt de poursuivre l'étude de codes de réseaux (et de classes latérales) non seulement pour aboutir à une meilleure compréhension de la philosophie du codage, mais également pour permettre d'envisager des solutions pratiques pour l'avenir.

**REFERENCES
ET
BIBLIOGRAPHIE**

REFERENCES ET BIBLIOGRAPHIE

A

[ABRA&STE 1972]

M. Abramowitz and I.A. Stegun (Ed.), "*Handbook of Mathematical Functions*", Washington, DC: National Bureau of Standards, 1972

[AND&MOH 1984]

J.B. Anderson and S. Mohan, "Sequential Coding Algorithms: A Survey and Cost Analysis", *IEEE Trans. Commun.*, COM 32, n.2, Feb., 1984, pp. 169-176

[AND *et al.* 1986]

J.B. Anderson, T. Aulin and C.E. Sundberg, "*Digital Phase Modulation*", New York: Plenum, 1986

[ANDR 1989]

F.T. Andrews, "The Heritage of Telegraphy", *IEEE Comm. Mag.*, 27, n.8, Aug., 1989, pp. 12-18

[ARI 1973]

S. Arimoto, "On the Converse to the Coding Theorem for Discrete Memoryless Channels", *IEEE Trans. Info. Theory*, IT 19, n.3, May, 1973, pp. 357-359

B

[BAH *et al.* 1974]

L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate", *IEEE Trans. Info. Theory*, IT 20, n.2, Mar., 1974, pp. 284-287

[BAR&SLO 1983]

E.S. Barnes and N.J.A. Sloane, "New lattice packings of spheres", *Can. J. Math.*, 35, n.1, 1983, pp. 117-130

[BAT&DEC 1976]

G. Battail and M. Decouvelaere, "Décodage par Répliques", *Ann. Télécomm.*, 31, n.11-12, nov-déc, 1976, pp. 387-404

[BAT 1983]

G. Battail, "Décodage pondéré optimal des codes linéaires en blocs I. Emploi simplifié du diagramme du treillis", *Ann. Télécomm.*, 38, n.11-12, nov-déc, 1983, pp. 443-459

[BAT 1984]

G. Battail, "L'évolution du codage vue par un ingénieur ou la dialectique de la fin et des moyens", [An engineer's point of view on the evolution of coding or the dialectic of the end and the means], *Trait. du Signal*, 1, n.2, 1984, pp. 87-96

[BAT 1989a]

G. Battail, "Construction Explicite de Bons Codes Longs", *Ann. Télécomm.*, 44, n.7-8, juillet-août, 1989, pp. 392-404

[BAT 1989b]

G. Battail, "Coding for the Gaussian Channel: The Promise of Weighted-Output

- Decoding", *Int. J. of Satellite Commun.*, 7, 1989, pp. 183-192
- [BAT 1989c]
G. Battail, "Building Long Codes by Combination of Simple ones, Thanks to Weighted-Output Decoding", *Proc. Int. Symp. on Signal, Systems, and Electronics, ISSSE'89*, Sept., Erlangen, FRG, 1989, pp. 634-637
- [BAT 1990]
G. Battail, "Communication en Présence de Bruit Gaussien", *L'Echo des Recherches*, n.39, 1er trimestre 1990, pp. 3-12
- [BAT et al. 1990]
G. Battail, H.M. de Oliveira and W. Zhang, "Coding and Modulation for the Gaussian Channel in the Absence or in the Presence of Fluctuations", *Int. Colloquium on Coding Theory and Applications, EUROCODE'90*, Nov., Udine, Italy, 1990 in: *Lecture Notes in Computer Science*, G. Cohen and P. Charpin (Eds.), EUROCODE'90 Proc., Berlin: Springer-Verlag, 1990, pp. 337-349
- [BAT 1991]
G. Battail, "Codage déterministe imitant le codage aléatoire", *13ème colloque GRETSI sur le traitement du signal et des images*, Juan-les-Pins, France, 16-20 sept. 1991, pp. 397-400
- [BAT&deO 1992]
G. Battail and H.M. de Oliveira, "Probabilité d'erreur du codage aléatoire avec décodage optimal sur le canal additif gaussien, affecté ou non de fluctuations", proposé aux *Ann. Télécomm.*
- [BAU&MIT 1989]
G. Baudoin et M.S. Mitrani, "TELSAT 19250: MODEM à 19200 bits/s sur Circuit Téléphonique", *Commutation & Transmission*, 2, 1989, pp. 73-84
- [BEA 1989]
N.C. Beaulieu, "A Simple Series for Personal Computer Computation of the Error Function $Q(\cdot)$ ", *IEEE Trans. Commun.*, COM 37, n.9, Sept., 1989, pp. 989-991
- [BE'E et al. 1989]
Y. Be'ery, B. Shahar and J. Snyders, "Fast Decoding of Leech Lattice", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 959-967
- [BEL et al. 1984]
W.H. Bellchambers, J. Francis, E. Hummel and R.L. Nickelson, "The International Telecommunication Union and Development of Worldwide Telecommunications", *IEEE Commun. Mag.*, 22, n.5, May, 1984, pp. 72-83
- [BEN et al. 1988]
S. Benedetto, M.A. Marsan, G. Albertengo and E. Giachin, "Combined Coding and Modulation: Theory and Applications", *IEEE Trans. Info. Theory*, IT 34, n.2, Mar., 1988, pp. 223-236
- [BEN et al. 1989]
S. Benedetto, M.A. Marsan, G. Masera and G. Olmo, "Encoded 16-PSK: A Study for Receiver Design", *IEEE Select. Areas on Commun.*, SAC 7, n.9, Dec., 1989, pp. 1381-1391
- [BER1 1969]
R.H. Beresford, "Statistical Geometry of Random Heaps of Equal Spheres", *Nature*, 224, Nov., 1969, pp. 550-553
- [BER et al. 1978]

E.R. Berlekamp, R.J. McEliece and H.C. Van Tilborg, "On the Inherent Intractability of Certain Coding Problems", *IEEE Trans. Info. Theory*, IT 24, n.3, May, 1978, pp. 384-386

[BER 1984]

E.R. Berlekamp, "Algebraic Coding Theory", New York: Aegean, 1984

[BER et al. 1987]

E.R. Berlekamp, R.E. Peile and S.P. Pope, "The Application of Error Control to Communications", *IEEE Commun. Mag.*, 25, n.4, Apr., 1987, pp. 44-57

[BIC et al. 1986]

J. Bic, D. Duponteil et J.C. Imbeaux, "Éléments de Communications Numériques", in: Collection Technique et Scientifique des Télécomm., Paris: BORDAS et CNET-ENST, 2 tomes, 1986

[BIG 1984]

E. Biglieri, "High-level Modulation and Coding for Nonlinear Satellite Channels", *IEEE Trans. Commun.*, COM 22, n.5, May, 1984, pp. 616-626

[BIG 1986]

E. Biglieri, "Ungerboeck Codes do not Shape the Signal Power Spectrum", *IEEE Trans. Info. Theory*, IT 32, n.4, Jun., 1986, pp. 595-596

[BIN 1988]

J.A.C. Bingham, "The Theory and Practice of MODEM design", New York: Wiley, 1988

[BLA 1971]

I.F. Blake, "The Leech Lattice as a Code for the Gaussian Channel", *Inform. Contr.*, 19, 1971, pp. 66-74

[BLA 1987]

R.E. Blahut, "Principles and Practice of Information Theory", Reading, MA: Addison-Wesley, 1987

[BOR&JOI 1988]

M. Borgne et M. Joindot, "Introduction aux Modulations Codées", Cours ENST 3ème année, Paris, 1988

[BÖR&SUN 1979]

P.O. Börjesson and C.E. Sundberg, "Simple Approximations of the Error Function $Q(x)$ for Communications Applications", *IEEE Trans. Commun.*, COM 27, n.3, Mar., 1979, pp. 639-643

[BOU 1964]

M. Bouix, "Les Fonctions Généralisées ou Distributions", Paris: Masson, 1964

C

[CAL et al. 1985]

A.R. Calderbank, J.E. Mazo and V.K. Wei, "Asymptotic Upper Bounds on the Minimum Distance of Trellis Codes", *IEEE Trans. Commun.*, COM 33, n.4, Apr., 1985, pp. 305-309

[CAL&SLO 1985]

A.R. Calderbank and N.J.A. Sloane, "Four-Dimensional Modulation with an 8-State Trellis Code", *AT&T Technical J.*, May-June, 1985, pp.1005-1018

[CAL&SLO 1987]

A.R. Calderbank and N.J.A. Sloane, "New Trellis Codes Based on Lattices and Cosets", *IEEE Trans. Info. Theory*, IT 33, n.2, Mar., 1987, pp. 177-195

[CAL 1990a]

A.R. Calderbank, "The Mathematics of MODEMs", *Mathematical intelligencer*, to appear.

[CAL 1990b]

A.R. Calderbank, "Binary Covering Codes and high speed Data Transmission", *Int. Colloquium on Coding Theory and Applications*, Invited Conf., EUROCODE'90, Nov., Udine, Italy, 1990 in: *Lecture Notes in Computer Science*, G. Cohen and P. Charpin (Eds.), EUROCODE'90 Proc., Berlin: Springer-Verlag, 1990, pp. 320-336

[CAL&OZA 1990]

A.R. Calderbank and L.H. Ozarow, "Non-Equiprobable Signaling on the Gaussian Channel", *IEEE Trans. Info. Theory*, IT 36, n.4, July, 1990, pp. 726-740

[CAM&GLA 1962]

C.N. Campopiano and B.G. Glazer, "A Coherent Digital Amplitude and Phase Modulation Scheme", *IRE Trans. Commun. Syst.*, CS 10, n.1, Mar., 1962, pp. 90-95

[CAS 1959]

J.W.S. Cassels, "An Introduction to the Geometry of Numbers", Gottingen/Berlin: Springer-Verlag, 1959

[CCITT 1989]

CCITT Recommendation de la Série V, "Famille de MODEMs à deux fils fonctionnant en duplex à des débits binaires allant jusqu'à 9600 bits/s pour usage sur le réseau téléphonique général avec commutation et sur les circuits loués de type téléphonique", [A Family of 2-wire, duplex Modems Operating at data Signalling Rates of up to 9,600 bits/sec for use on the general Switched Telephone network and leased telephone-type Circuits], *Communication de Données sur le Réseau Téléphonique*, in: Livre Bleu, tome VIII, fasc.VIII.1, Geneva, Switzerland, 1989

[CCITT 1989]

CCITT Recommendation de la Série V, "MODEM à 14400 bits/s Normalisé pour Usage sur Circuits Loués à quatre fils poste à poste, de Type Téléphonique", [A 14,400 bits/sec Modem Standardized for use on point-to-point 4-wire leased telephone-type Circuits], *Communication de Données sur le Réseau Téléphonique*, in: Livre Bleu, tome VIII, fasc.VIII.1, Geneva, Switzerland, 1989

[CHA 1988]

L.F. Chang, "Block-Coded-Modulation for digital radio Communications Systems", *Proc. Int. Conf. Commun.*, ICC'88, Philadelphia, PA, Jun., 1988, pp. 24.5.1-24.5.6

[CHE 1989]

K-M. Cheung, "More on the Decoder Error Probability for Reed-Solomon Codes", *IEEE Trans. Info. Theory*, IT 35, n.4, July, 1989, pp. 895-900

[CHE 1990]

K-M. Cheung, "Identities and Approximations for the Weight Distribution of Q-ary Codes", *IEEE Trans. Info. Theory*, IT 36, n.5, Sept., 1990, pp. 1149-1153

[CHO&SAR 1989]

A. Chouly and H. Sari, "Block Coded Modulation Schemes Based on Two-Step Partitioning of the Signal Alphabet", *12ème colloque GRETSI sur le traitement du signal et des images*, Juan-les-Pins, France, 12-16 juin 1989, pp.275-278

[CHO&SAR 1990]

- A. Chouly and H. Sari, "Design and Performance of Block-Coded Modulation for Digital Microwave Radio Systems", *IEEE Trans. Commun.*, COM 38, n.5, May, 1990, pp. 727-733
- [CLA&CAI 1981]
G.C. Clark and J.B. Cain, "Error-Correction for Digital Communications", New York: Plenum, 1981
- [COF&GOO 1990]
J.T. Coffey and R.M. Goodman, "Any Code of which we cannot Think is Good", *IEEE Trans. Info. Theory*, IT 36, n.6, Nov., 1990, pp. 1453-1461
- [CON&SLO 1982a]
J.H. Conway and N.J.A. Sloane, "Voronoi Regions of Lattices, Second Moment of Polytopes and Quantization", *IEEE Trans. Info. Theory*, IT 28, n.2, Mar., 1982, pp. 211-226
- [CON&SLO 1982b]
J.H. Conway and N.J.A. Sloane, "Fast Quantizing and Decoding Algorithms for Lattice Quantizers and Codes", *IEEE Trans. Info. Theory*, IT 28, n.2, Mar., 1982, pp. 227-232
- [CON&SLO 1986]
J.H. Conway and N.J.A. Sloane, "Soft Decoding Techniques for Codes and Lattices, Including the Golay Code and the Leech Lattice", *IEEE Trans. Info. Theory*, IT 32, n.1, Jan., 1986, pp. 41-50
- [CON&SLO 1988]
J.H. Conway and N.J.A. Sloane, "Sphere Packings, Lattices and Groups", New York: Springer-Verlag, 1988
- [CUP&GIT 1989]
R.L. Cupo and R.D. Gitlin, "Adaptive Carrier Recovery Systems for Digital Data Communications Receivers", *IEEE J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1328-1339
- [CUS 1984]
E.L. Cusack, "Error Control Codes for QAM Signaling", *Elect. Lett.*, 20, 1984, pp. 62-63

D

- [deB 1975]
R. de Buda, "The Upper Error Bound of a New Near-Optimal Code", *IEEE Trans. info. Theory*, IT 21, n.4, July, 1975, pp. 441-445
- [deB 1982]
R. de Buda, "About the Reliability Exponent of Lattice Codes", presented at *IEEE Int. Symp. Info. Theory*, ISIT'82, Les Arcs, France, Jun. 1982, p. 76
- [deB 1989]
R. de Buda, "Some Optimal Codes have Structure", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 893-899
- [deO 1986]
H.M. de Oliveira, "Empirical Bayes Adaptive Decoding for Sources with Unknown Distribution", presented at *IEEE Int. Symp. Info. Theory*, ISIT'86, Ann Arbor, USA, Oct. 1986, p. 108

[deO&BAT 1990]

H.M. de Oliveira and G. Battail, "A Capacity Theorem for Lattice Codes on Gaussian Channels", *Proc. IEEE/STB Int. Telecomm. Symp.*, ITS'90, Sept., Rio de Janeiro, Brazil, 1990, pp. 1.2.1-1.2.5

[deO&BAT 1991]

H.M. de Oliveira and G. Battail, "On Generalized Constellations and the Opportunistic Secondary Channel", submitted to *Ann. Télécomm.*

[deO&BAT 1991]

H.M. de Oliveira and G. Battail, "The Random Coded-Modulation: Performance and Euclidean Distance Spectrum Evaluation", *Proc. of IEEE Int. Symp. Info. Theory*, ISIT'91, Budapest, Hungary, Jun. 1991, p. 2

[deO et al. 1992]

H.M. de Oliveira, W. Zhang and G. Battail, "Performance of Lattice Codes on Gaussian Channels", proposé aux *Ann. Télécomm.*

[DIG 1976]

A. Digeon, "Codage Convolutionnel de l'Information", *Ann. Télécomm.*, 31, n.5-6, 1976, pp. 190-198

[DIG 1977]

A. Digeon, "On Improving Bit Error Probability of QPSK and 4-level Amplitude Modulation Systems by Convolutional Coding", *IEEE Trans. Commun.*, COM 25, n.10, Oct., 1977, pp. 1238-1239

[DIV et al. 1987]

D. Divsalar, M.K. Simon and J.H. Yuen, "Trellis Coding with Asymmetric Modulations", *IEEE Trans. Commun.*, COM 35, n.2, Feb., 1987, pp. 130-141

[DIV&SIM 1987]

D. Divsalar and M.K. Simon, "Generalized Multiple Trellis Code Modulation (MTCM)", *Proc. Int. Conf. Commun.*, ICC'87, Jun., Seattle, WA, 1987, pp. 20.3.1-20.3.7

[DIV&SIM 1988a]

D. Divsalar and M.K. Simon, "Multiple Trellis Coded Modulation (MTCM)", *IEEE Trans. Commun.*, COM 36, n.4, Apr., 1988, pp. 410-419

[DIV&SIM 1988b]

D. Divsalar and M.K. Simon, "The Design of Trellis Coded MPSK for Fading Channels: Performance Criteria", *IEEE Trans. Commun.*, COM 36, n.9, Sept., 1988, pp. 1004-1012

[DIV&SIM 1988c]

D. Divsalar and M.K. Simon, "The Design of Trellis Coded MPSK for Fading Channels: Set Partitioning for Optimum Code Design", *IEEE Trans. Commun.*, COM 36, n.9, Sept., 1988, pp. 1013-1021

[DU&VUC 1991]

J. Du and B. Vucetic, "New 16-QAM Trellis Codes for Fading Channels", *Elect. Lett.*, 27, n.12, 1991, pp. 1009-1010

E

[EDB 1987]

F. Edbauer, "Coded 8-DPSK Modulation with Differentially Coherent Detection:

An Efficient Modulation Scheme for Fading Channels", *Proc. Global Telecommun. Conf., GLOBECOM'87*, Nov., Tokyo, Japan, 1987, pp. 42.21-42.24

[EDB 1989a]

F. Edbauer, "Concatenated Trellis-Coded Differential 8-PSK Modulation for Rayleigh Channels", *Proc. Int. Symp. on Signal, Systems, and Electronics, ISSSE'89*, Sept., Erlangen, FRG, 1989, pp. 268-271

[EDB 1989b]

F. Edbauer, "Performance of Interleaved Trellis-Coded Differential 8-PSK Modulation over Fading Channels", *IEEE J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1340-1346

[ELI 1955]

P. Elias, "Coding for Noisy Channels", *IRE Conv. Rec.*, pt.4, pp. 37-46, Mar., 1955 in: Key papers in "The Development of Info. Theory", D. Slepian (Ed.), New York: IEEE Press, 1973

[EYU&QUR 1989]

M.V. Eyuboglu and S.U.H. Qureshi, "Reduced-State Sequence Estimation for Coded Modulation on Intersymbol Interference Channels", *IEEE J. Select. Areas Commun.*, SAC 6, n.6, Aug., 1989, pp. 989-995

F

[FOR 1970]

G.D. Forney Jr., "Coding and its Application in Space Communications", *IEEE Spectrum*, 7, Jun., 1970, pp. 47-58

[FOR 1973]

G.D. Forney Jr., "The Viterbi Algorithm", *Proc. IEEE*, 61, Mar., 1973, pp. 268-278

[FOR et al. 1984]

G.D. Forney Jr., R.G. Gallager, G.R. Lang, F.M. Longstaff and S.U. Qureshi, "Efficient Modulation for Band-limited Channels", *IEEE J. Select. Areas Commun.*, SAC 2, n.5, Sept., 1984, pp. 632-646

[FOR 1988a]

G.D. Forney Jr., (D.G. Fink Prize Award 1990) "Coset codes- part I: Introduction and Geometrical Classification", *IEEE Trans. Info. Theory*, IT 34, n.5, Sept., 1988, pp. 1123-1151

[FOR 1988b]

G.D. Forney Jr., (D.G. Fink Prize Award 1990) "Coset codes- part II: Binary Lattices and Related Codes", *IEEE Trans. Info. Theory*, IT 34, n.5, Sept., 1988, pp. 1152-1187

[FOR&WEI 1989]

G.D. Forney Jr. and L-F. Wei, "Multidimensional Constellation part I: Introduction, Figures of Merit, and Generalized Cross Constellations", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 877-892

[FOR 1989a]

G.D. Forney Jr., "Multidimensional Constellations- part II: Voronoi Constellations", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 941-958

[FOR 1989b]

G.D. Forney Jr., "A Bounded-Distance Decoding Algorithm for the Leech Lattice",

IEEE Trans. Info. Theory, IT 35, n.4, July/Aug., 1989, pp. 906-909

[FOR&CAL 1989]

G.D. Forney Jr. and A.R. Calderbank, "Coset Codes for Partial Response Channel; or Coset Codes with Spectral Nulls", *IEEE Trans. Info. Theory*, IT 35, n.5, Sept., 1989, pp. 925-943

[FOR 1990]

G.D. Forney Jr., Review of "Sphere Packings, Lattices and Groups" by J.H. Conway/N.J.A. Sloane, *IEEE Trans. Info. Theory*, (Book Reviews), IT 36, n.4, July, 1990, pp.955-956

[FOS *et al.* 1974]

G.J. Foschini, R.D. Gitlin and S.B. Weinstein, "Optimization of 2-Dimensional Signal Constellations in Presence of Gaussian noise", *IEEE Trans. Commun.*, COM 22, n.1, Jan., 1974, pp. 28-37

G

[GAL 1968]

R.G. Gallager, "*Information Theory and Reliable Communication*", New York: Wiley, 1968

[GAL 1965]

R.G. Gallager, "A simple Derivation of the Coding Theory and some Applications", *IEEE Trans. Info. Theory*, IT 11, Jan., 1965, pp. 3-18

[GER&LAW 1984]

A. Gersho and V.B. Lawrence, "Multidimensional Signal Constellations for Voice Band data Transmission", *IEEE J. Select. Areas Commun.*, SAC 2, n.5, Sept., 1984, pp. 687-702

[GIT 1984]

R.D. Gitlin, "Recent Trends in Voiceband data Transmission Techniques", presented at *Int. Conf. Commun.*, ICC'84, Amsterdam, The Netherlands, Jun., 1, 1984, p. 12.1

H

[HAG 1989]

J. Hagenauer, "Coded Phase Modulation - An Overview", *Proc. Int. Symp. on Signals, Systems, and Electronics*, ISSSE'89, Sept., Erlangen, FRG, 1989, pp. 255-260

[HAJ 1990]

D. Hajela, "On Computing the Minimum Distance for Faster than Nyquist Signaling", *IEEE Trans. Info. Theory*, IT 36, n.2, Mar., 1990, pp. 289-295

[HAK 1977]

S.J. Hakowitz, "*Computational Probability and Simulation*", Reading, MA: Addison-Wesley, 1977

[HAM 1980]

R.W. Hamming, "*Coding and Information Theory*", Englewood Cliffs NJ:Prentice-Hall, 1980

[HAR 1928]

R.V.L. Hartley, "Transmission of Information", *Bell Syst. Tech. J.*, 7, 1928, pp.

535-563

[HEL&STI 1973]

H.J. Helgert and R.D. Stinaff, "Minimum-Distance Bounds for Binary Linear Codes", *IEEE Trans. Info. Theory*, IT 19, n.3, May, 1973, pp. 344-356

[HEM 1989]

F. Hemmati, "Closest Coset Decoding of $[u|u+v]$ Codes", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 982-988

[HER&POL 1991]

H. Herzberg and G. Polyrev, "The use of Random Coset Coding for Bounding the Probability of error in Coded Modulation Schemes", *Proc. 17th Conf. of IEEE in Israel*, Israel, 1991.

[HIR 1980]

B. Hirosaki, "An Analysis of Automatic Equalizers for Orthogonal Multiplexed QAM Systems", *IEEE Trans. Info. Theory*, COM 28, n.1, Jan., 1980, pp. 73-83

[HIR 1981]

B. Hirosaki, "An Orthogonally Multiplexed QAM System Using Discrete Fourier Transform", *IEEE Trans. Info. Theory*, COM 29, n.7, July, 1981, pp. 982-989

[HIR 1985]

B. Hirosaki, "A 19.2 Kbps Voice-Band data Modem Based on Orthogonally Multiplexed QAM Techniques", *Proc. Int. Conf. Commun.*, ICC'85, Jun., Chicago, 1985, pp. 21.1.1-21.1.4

[HIRS 1975]

P. Hirschler, "Théorie des communications: Les limites du possible", *Ann. des Mines*, nov., 1975, pp. 39-50

[HOL 1977]

J.L. Holsinger, "Where are Modems Going?", *IEEE Commun. Mag.*, Sept., 1977, pp. 3-5,10

[HON 1986]

M.L. Honig, "Optimization of Trellis Codes with Multilevel Amplitude Modulation with Respect to an Error Probability Criterion", *IEEE Trans. Commun.*, COM 34, 1986, pp. 821-825

[HON 1988]

M.L. Honig, "On Constructing Embedded Multilevel Trellis Codes", *IEEE Trans. Commun.*, COM 36, n.2, Feb., 1988, pp. 218-221

[HUB 1989]

K. Huber, "Combined Coding and Modulation Using Block Codes", *Elect. Lett.*, 25, n.17, Aug., 1989, pp. 1130-1131

[HUG 1991]

B. Hughes, "On the Error Probability of Signals in Additive White Gaussian Noise", *IEEE Trans. Info. Theory*, IT 37, n.1, Jan., 1991, pp. 151-155

I

[ISA 1988]

M. Isaksson, "A Class of Block Codes with Expanded Signal-Sets of PSK Modulation Properties and Information Retrieval", *Royal Inst. of Technol. Telecomm. Theory*, report n. TRITA-TTT-8806, Aug., 1988, Stockholm, Sweden

[ISA 1989]

M. Isaksson, "Decoding of a Class of Block Codes with Expanded Signal Sets for PSK-Modulation", *Proc. Int. Symp. on Signal, Systems, and Electronics, ISSSE'89*, Sept., Erlangen, FRG, 1989, pp. 255-260

J

[JAM&LE-N 1991]

S.H. Jamali and T. Le-Ngoc, "A New 4-State 8PSK TCM Scheme for Fast Fading, Shadowed Mobile Radio Channels", *IEEE Trans. Veh. Tech.*, 40, n.1, Feb., 1991, pp. 216-222

[JEL 1968]

F. Jelinek, "Probabilistic Information Theory", New York: McGraw-Hill, 1968

[JER 1977]

A.J. Jerri, "The Shannon Sampling Theorem - its various extensions and applications: A Tutorial Review", *Proc. IEEE*, 65, n.11, 1977, pp. 1565-1596.

K

[KAB&PAS 1975]

P. Kabal and S. Pasupathy, "Partial-Response Signaling", *IEEE Trans. Commun.*, COM 23, n.9, Sept., 1975, pp. 921-934

[KAR et al. 1991]

G. Karam, A. Chouly and S. Moridi, "Analyse des Performances du Codage Multiniveau pour Modulations à grande nombre d'états", *Ann. Télécomm.*, 46, n.5-6, 1991, pp. 310-318

[KAS&DU 1988]

M. Kasahara and J. Du, "Coded Modulation System with Block Codes over GF(2) and GF(2**m)", presented at *IEEE Int. Symp. Info. Theory, ISIT'88*, Kobe, Jun. 1988, Japan, p. 188

[KAW-KAL 1989]

G. Kawas-Kaleh, "Joint Decoding and Phase Estimation via the Expectation-Maximization Algorithm", presented at *IEEE Int. Symp. Info. Theory, ISIT'90*, San Diego, USA, Jan. 1990

[KNU 1975]

D.E. Knuth, "The Art of Computer Programming", Vol 2 (Seminumerical algorithms), Reading, MA: Addison-Wesley, 1975

[KOR&KOR 1968]

G.A. Korn and T.M. Korn, "Mathematical Handbook for Scientists and Engineers", New York: McGraw-Hill, 2nd.Ed, 1968

[KSC et al. 1989]

F.R. Kschischang, P.G. de Buda and S. Pasupathy, "Block Coset Codes for M-ary Phase Shift Keying", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 900-913

L

[LAN&LON 1989]

G.R. Lang and F.M. Longstaff, "A Leech Lattice MODEM", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 968-973

[LAZ&SEN 1990]

D.E. Lazić and V. Šenk, "A Simple Geometrical Method for Bounding the Error Exponent for Specific Families of Channel Block Codes- part I: Cutoff Rate - Lower Bound", submitted to *IEEE Trans. Info. Theory*

[LEECH 1967]

J. Leech, "Notes on Sphere Packings", *Canad. J. Math.*, 19, 1967, pp. 251-267

[LEECH&SLO 1971]

J. Leech and N.J.A. Sloane, "Sphere Packings and Error-Correcting Codes", *Canad. J. Math.*, 23, 1971, pp. 718-745

[LEEU *et al.* 1990]

K. Leeuwin-Boullé, J.C. Belfiore and G. Kawas-Kaleh, "Chernoff Bound of Trellis Coded Modulation over Correlated Rayleigh Channel", *Proc. Int. Colloquium on Coding Theory and Applications*, EUROCODE'90, Nov., Udine, Italy, 1990, in: *Lecture Notes in Computer Science*, G. Cohen and P. Charpin (Eds.), EUROCODE'90 Proceedings, Berlin: Springer-Verlag, 1990, pp. 364-376

[LEEU&BEL 1991]

K. Leeuwin-Boullé and J.C. Belfiore, "Construction of Constellations designed for the Rayleigh Channel", *13ème Colloque GRETSI sur le traitement des signal et des images*, Juan-les-Pins, France, 16-20 sept. 1991, pp. 437-440

[LEE 1986]

P.J. Lee, "Computation of the bit Error Rate of Coherent M-ary PSK with Gray Code bit Mapping", *IEEE Trans. Commun.*, COM 34, n.5, May, 1986, pp. 488-491

[LEE&FAR 1991]

L.H.C. Lee and P.G. Farrell, "Performance of Phase-invariant Trellis Coding for Differentially-encoded M-PSK Signals on Rayleigh Fading Channels", *Elec. Lett.*, 27, n.12, 1991, pp. 1051-1052

[LIN&COS 1983]

S. Lin and D.J. Costello Jr., "Error Control Coding: Fundamentals and Applications", Englewood Cliffs, NJ: Prentice-Hall, 1983

[LIU *et al.* 1990]

Y-J. Liu, I. Oka and E. Biglieri, "Error Probability for Digital Transmission Over Nonlinear Channels with Application to TCM", *IEEE Trans. Info. Theory*, IT 36, n.5, Sept., 1990, pp. 1101-1110

[LOR *et al.* 1989]

F. Lorenzelli, L. Testa, M. Visintin, E. Biglieri and M. Pent, "Clock-aided Carrier Recovery in Trellis-Coded PSK", *IEEE J. Select. Areas Commun. Theory*, SAC 7, n.9, Dec., 1989, pp.1307-1317

M

[MacW&SLO 1977]

F.J. MacWilliams and N.J.A. Sloane, "The Theory of Error-Correcting Codes", Amsterdam, The Netherlands: North-Holland, 1977

[MAK&FEH 1988]

D. Makrakis and K. Feher, "On Optimal Detection of non-coherent Trellis Coded Modulation Signal: Differential Detection", *38th IEEE Veh. Tech. Conf.*, Philadelphia, PA, Jun., 1988, pp. 1-5

[MAR *et al.* 1983]

M.A. Marsan, G. Albertengo, S. Benedetto and D. Del Corso, "A 9600 bps Microprocessor Modem Prototype", *Alta Frequenza*, LII, n.4, 1983, pp. 224-255

[MAR *et al.* 1984]

M.A. Marsan, G. Albertengo, S. Benedetto and E. Giachin, "Theoretical Aspects and Performance Evaluation of Channel Encoding with Redundant Modulation Alphabets", *Proc. Global Telecommun. Conf.*, GLOBECOM'84, Nov., Atlanta, GA, 1984, pp. 20.5.1-20.5.5

[MAR *et al.* 1987]

M.A. Marsan, G. Albertengo and S. Benedetto, "Performance Evaluation of Combined Coding and Modulation Schemes for Nonlinear Channels", *IEEE Trans. Commun.*, COM 35, n.9, Sept., 1987, pp. 969-972

[MAS 1974]

J.L. Massey, "Coding and Modulation in Digital Communications", *Proc. Int. Zurich Seminar on Digital Communications*, Zurich, Switzerland, Mar., 1974, pp. E2(1)-(4)

[MAS 1989]

J.L. Massey, "A Short Introduction to Coding Theory and Practice", *Proc. Int. Symp. of Signals, Systems, and Electronics*, ISSSE'89, Sept., Erlangen, FRG, 1989, pp. 629-633

[MAZ&LAN 1988]

J.E. Mazo and H.J. Landau, "On the Minimum Distance Problem for Faster-than-Nyquist Signaling", *IEEE Trans. Info. Theory*, IT 34, n.6, Nov., 1988, pp. 1420-1427

[MAZ 1975]

J.E. Mazo, "Faster-than-Nyquist Signaling", *Bell Syst. Tech. J.*, 54, n.8, Oct., 1975, pp. 1451-1462

[McL *et al.* 1987]

P.J. McLane, P.H. Witke, P.K.M. Ho and C. Loo, "PSK and DPSK Trellis Codes for Fast Fading, Shadowed Mobile Satellite Communication Channels", *Proc. Int. Conf. Commun.*, ICC'87, Jun., Seattle, WA, 1987, pp. 21.1.1-21.1.6

[McL *et al.* 1988]

P.J. McLane, P.H. Witke, P.K.-M Ho and C. Loo, "PSK and DPSK Trellis Codes for Fast Fading, Shadowed Mobile Satellite Communication Channels", *IEEE Trans. Commun.*, COM 26, n.11, Nov., 1988, pp. 1242-1246

[MIL 1984]

S. Millman, (Ed.), "A History of Engineering and Science in the Bell Systems: Communications Science (1925-1980)", Indianapolis:AT&T Tech., 1984

[MOH&LOD 1989]

M.L. Moher and J.H. Lodge, "TCMP- A Modulation and Coding Strategy for Rician Fading Channels", *IEEE J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1347-135x

[MUD 1990]

D.J. Muder, "Minimal Trellises for Block Codes", *IEEE Trans. Info. Theory*, IT 34,

n.5, Sept., 1990, pp. 1049-1053

N

[NIL 1991]

J. Nilsson, "A Geometric Analysis of Algebraic Soft Decoding", PhD Thesis, Linköping University, Sweden, Feb., 1991.

[NYM *et al.* 1982]

H. Nyman, F. Salcedo, J.T.L. Sharpe and C.W. Tarry, "Evolution des Modems pour la Transmission de Données", *Revue des Télécomm.*, 57, n.3, 1982, pp. 187-194

[NYQ 1924]

H. Nyquist, "Certain Factors Affecting Telegraph Speed", *Bell Syst. Tech. J.*, Apr., 1924, pp. 324-346

[NYQ 1928]

H. Nyquist, "Certain Topics in Telegraph Transmission Theory", *Trans. of the A.I.E.E.*, 47, 1928, pp. 617-644

O

[OLI *et al.* 1948]

B.M. Oliver, J.R. Pierce and C.E. Shannon, *Proc. IRE*, 36, Nov., 1948, pp. 1384-1331

P

[PAD&WOL 1986]

R. Padovani and J.K. Wolf, "Coded Phase/Frequency Modulation", *IEEE Trans. Commun.*, COM 34, n.5, May, 1986, pp. 446-453

[PAH&HOL 1988]

K. Pahlavan and J.L. Holsinger, "Voice-Band data Communication Modems - A Historical Review 1919-1988", *IEEE Commun. Mag.*, 26, n.1, Jan., 1988, pp. 16-27

[PAU&SCH 1987]

C.K. Pauw and D.L. Schilling, "Probability of Error for M-ary DPSK on Rayleigh Fading Channel", *Proc. Global Telecommun. Conf.*, GLOBECOM'87, Nov., Tokyo, Japan, 1987, pp. 25.4.1-25.4.4

[PIE 1967]

J.N. Pierce, "Limit Distribution of the Minimum Distance of Random Linear Codes", *IEEE Trans. Info. Theory*, IT 13, n.4, Oct., 1967, pp. 595-599

[PIE *et al.* 1990]

J.S. Pietrobon, R.H. Deng, A. Lafanechere, G. Ungerboeck and D.J. Costello Jr., "Trellis-Coded Multidimensional Phase Modulation", *IEEE Trans. Info. Theory*, IT 36, n.1, Jan., 1990, pp. 63-89

[POR 1991]

J. Portugheis, "On block-coded multidimensional M-PSK modulation", *Proc. Symp. Info. Theory Appl.*, SITA'91, Ibusuki, Japan, Dec., 1991

[POT&TAY 1987]

G.J. Pottie and D.P. Taylor, "An Approach to Ungerboeck Coding for Rectangular

Signal Sets", *IEEE Trans. Info. Theory*, IT 33, n.2, Mar., 1987, pp. 285-290

[POT&TAY 1989]

G.J. Pottie and D.P. Taylor, "A Comparison of Reduced Complexity Decoding Algorithms for Trellis Codes", *IEEE J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1369-1380

[PRO 1989]

J.G. Proakis, "Digital Communications", New York: McGraw-Hill, 2nd Ed., 1989

R

[RAM 1970]

J.L. Ramsey, "Realization of Optimum Interleavers", *IEEE Trans. Info. Theory*, IT 16, n.3, May, 1970, pp. 338-345

[REE&SOL 1960]

I.S. Reed and G. Solomon, "Polynomial Codes over Certain Finite Fields", *J. Soc. Indust. Appl. Math.*, 8, n.2, Jun., 1960, pp. 300-304

[ROB 1969]

G.H. Robertson, "Computation of the Noncentral Chi-Square Distribution", *Bell Syst. Tech. J.*, 48, n.1, Jan., 1969, pp. 201-207

[ROU&COS 1987]

M. Rouanne and D.J. Costello Jr., "A Lower Bound on the Minimum Euclidean Distance of Trellis Codes", in: *Coding Theory and Applications*, G. Cohen /P. Godlewski (Eds.), Springer Verlag, 1987, pp. 130-140

[ROU 1988]

M. Rouanne, "Distance Bounds and Construction Algorithms for Trellis Codes", PhD thesis, Univ. of Notre Dame, Indiana, Apr., 1988

[ROU&COS 1989]

M. Rouanne and D.J. Costello Jr., "An Algorithm for Computing the Distance Spectrum of Trellis Codes", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 929-940

[ROU&COS 1990]

M. Rouanne and D.J. Costello Jr., "A Lower Bound on the Minimum Euclidean Distance of Trellis-Coded Modulation Schemes", *IEEE Trans. Info. Theory*, IT 34, n.5, Sept., 1990, pp. 1011-1020

S

[SAY 1986]

S.I. Sayegh, "A Class of Optimum Block Codes in Signal Space", *IEEE Trans. Commun.*, COM 34, n.10, Oct., 1986, pp. 1043-1045

[SCH&COS 1989]

C. Schlegel and D.J. Costello Jr., "Bandwidth Efficient Coding for Fading Channels: Code Construction and Performance Analysis", *IEEE J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1356-1368

[SEC&deB 1989]

N.P. Secord and R. de Buda, "A Two-Stage Sequential Demodulator for the Gosset

- Lattice", *IEEE J. Select. Areas Commun.*, SAC 7, n.6, Aug., 1989, pp. 974-981
- [SHA 1948a]
C.E. Shannon, "A Mathematical Theory of Communication", *Bell Syst. Tech. J.*, **45**, July, 1948, pp. 379-423
- [SHA 1948b]
C.E. Shannon, "A Mathematical Theory of Communication", *Bell Syst. Tech. J.*, **17**, Oct., 1948, pp. 623-656
- [SHA 1949]
C.E. Shannon, "Communication in Presence of Noise", *Proc. IRE*, **37**, Jan., 1949, pp. 10-21
- [SHA 1959]
C.E. Shannon, "Probability of Error for Optimal Codes in a Gaussian Channel", *Bell Syst. Tech. J.*, **38**, n.3, May, 1959, pp. 611-656
- [SLO 1980]
N.J.A. Sloane, "A Note of the Leech Lattice as a Code for the Gaussian Channel", *Inform. Contr.*, **46**, 1980, pp. 270-272
- [SLO 1981]
N.J.A. Sloane, "Tables of Sphere Packings and Spherical Codes", *IEEE Trans. Info. Theory*, IT 27, n.3, 1981, pp. 327-338
- [SLO 1984]
N.J.A. Sloane, "The Packing of Spheres", *Scientific American*, **250**, Jan., 1984, pp. 92-101 See also "Les Empilements de Sphères", in: *Les Mathématiques Aujourd'hui*, Paris: Pour la Science, 1984, pp. 53-63
- [SIM&DIV 1987]
M.K. Simon and D. Divsalar, "The Performance of Trellis Coded Multilevel DPSK on a Fading Mobile Satellite Channel", *JPL Publication (NASA)*, MSAT-X report n.144, Jun., 1987, Pasadena, CA
- [STE 1987]
M. Stein, "*Les MODEMS pour Transmission de Données*", Paris: Masson et CNET-ENST, 1987
- [STU 1984]
F.L.H.M. Stumpers, "The History Development, and Future of Telecommunications in Europe", *IEEE Commun. Mag.*, **22**, n.5, May, 1984, pp. 84-95
- [SUN 1986]
C.E. Sundberg, "Continuous Phase Modulation", *IEEE Commun. Mag.*, **24**, n.4, Apr., 1986, pp. 25-38

T

- [TAN *et al.* 1987]
Y. Tanaka, S. Unagami, T. Kinoshita, T. Kaku and Y. Arai, "A 19.2 kb/s High-Speed Voiceband data Modem Based on Eight-Dimensional Trellis Coding", *Proc. Global Telecommun. Conf.*, GLOBECOM'87, Nov., 1987, Tokyo Japan, pp. 36.1.1-36.1.5
- [TES 1990]
M. Testard, "Modulations Codées pour les faisceaux Hertiens à grand débit", presented at *Int. Colloquium on Coding Theory and Applications*, EUROCODE'90,

Nov., Udine, Italy, 1990

[THA 1984]

H.K. Thapar, "Real-time Application of Trellis Coding to High-speed Voice-Band data Transmission", *IEEE J. Select. Areas Commun.*, SAC 2, n.5, Sept., 1984, pp. 648-658

[THOM *et al.* 1974]

C.M. Thomas, M.Y. Weidner and S.H. Durrani, "Digital Amplitude-Phase Keying with M-ary Alphabets", *IEEE Trans. Commun.*, COM 22, n.2, Feb., 1974, pp. 168-180

[THOMP 1983]

T.M. Thompson, "From Error-Correcting Codes Through Sphere Packings to Simple Groups", USA: The Math. Association of America, 1983

[TOR *et al.* 1991]

J. Torck, S. Bellini and G. Kawas-Kaleh, "Décodage de Viterbi et estimation de phase conjoints", *13ème colloque GRETSI sur le traitement du signal et des images*, Juan-les-Pins, France, 16-20 sept. 1991, pp. 469-471

[TRE 1989]

S.A. Tretter, "An Eight-Dimensional 64-State Trellis Code for Transmitting 4 Bits per 2-D Symbol", *J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1392-1395

U

[UNG&CSA 1976]

G. Ungerboeck and I. Csajka, "On Improving Data Line Performance by Increasing the Channel Alphabet and Introducing Sequence Coding", presented at *IEEE Int. Symp. Info. Theory, ISIT'76*, Ronneby, Sweden, Jun. 1976, p. 53

[UNG 1982]

G. Ungerboeck, (1984 Info. Theory Society Paper Award) "Channel Coding with Multilevel/phase Signals", *IEEE Trans. Info. Theory*, IT 28, n.1, Jan., 1982, pp. 55-67

[UNG 1987a]

G. Ungerboeck, "Trellis-coded Modulation with Redundant Signal Sets. Part I: Introduction", *IEEE Commun. Mag.*, 25, n.2, Feb., 1987, pp. 12-21

[UNG 1987b]

G. Ungerboeck, "Trellis-coded Modulation with Redundant Signal Sets. Part II: State of the Art", *IEEE Commun. Mag.*, 25, n.2, Feb., 1987, pp. 12-21

V

[VER 1987]

T. Verhoeff, "An Updated Table of Minimum-Distance Bounds for Binary Linear Codes", *IEEE Trans. Info. Theory*, IT 33, n.5, Sept., 1987, pp. 665-680

[VIT 1967]

A.J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimal Decoding Algorithm", *IEEE Trans. Info. Theory*, IT 13, n.2, Apr., 1967, pp. 260-269

[VIT 1971]

A.J. Viterbi, "Convolutional Codes and their Performance in Communications Systems", *IEEE Trans. Commun. Technol.*, COM 19, n.5, Oct., 1971, pp. 751-772

[VIT&OMU 1982]

A.J. Viterbi and J.K. Omura, "*Principes des Communications Numériques*" [Principles of Digital Communication and Coding], in: Collection technique et scientifique des Télécommunications, Paris:Dunod, Bordas et CNET-ENST, 1982

[VIT et al. 1989]

A.J. Viterbi, J.K. Wolf, E. Zehavi and R. Padovani, "A Pragmatic Approach to Trellis-coded Modulation", *IEEE Commun. Mag.*, 27, n.7, July, 1989, pp. 11-19

[VO 1991]

C. Voß, "A Family of Binary Codes of Geometric Goppa Codes and the Gilbert-Varshamov bound", in: Lecture Notes in Computer Science, G. Cohen and P. Charpin (Eds.), EUROCODE'90 Proceedings, Berlin: Springer-Verlag, 1990, pp. 150-157

[VU et al. 1990]

B. Vucetic, L. Zhang and G. Kaschatrion, "Construction of Block Modulation Codes over Rings for Fading Channels", *Elect. Lett.*, 26, n.24, Nov., 1990, pp. 2020-2022

W

[WEA&SHA 1949]

W. Weaver and C.E. Shannon, "*Théorie Mathématique de la Communication*", [The Mathematical Theory of Communication, Univ. of Illinois], Paris: C.E.P.L., 1975

[WEI 1984a]

L-F. Wei, "Rotationally Invariant Convolutional Channel Coding with Expanded Signal Space. Part I: 180 degrees", *IEEE J. Select. Areas Commun.*, SAC 2, n.5, Sept., 1984, pp. 659-672

[WEI 1984b]

L-F. Wei, "Rotationally Invariant Convolutional Channel Coding with Expanded Signal Space. Part II: nonlinear codes", *IEEE J. Select. Areas Commun.*, SAC 2, n.5, Sept., 1984, pp. 672-686

[WEI 1987]

L-F. Wei, (1989 Info. Theory Society Paper Award) "Trellis-Coded Modulation with Multidimensional Constellations", *IEEE Trans. Info. Theory*, IT 33, n.4, July, 1987, pp. 483-501

[WEI 1989]

L-F. Wei, "Rotationally Invariant Trellis-Coded Modulations with Multidimensional M-PSK", *IEEE J. Select. Areas Commun.*, SAC 7, n.9, Dec., 1989, pp. 1281-1295

[WIL et al. 1984]

S.G. Wilson, H.A. Sleeper II, P.J. Schottler and M.T. Lyons, "Rate 3/4 Convolutional Coding of 16 PSK Code Design and Performance Study", *IEEE Trans. Commun.*, COM 32, n.12, Dec., 1984, pp. 1308-1315

[WIL et al. 1984]

S.G. Wilson, H.A. Sleeper II and N.K. Srinath, "Four-Dimensional Modulation and Coding: An Alternative to Frequency-reuse", *Proc. Int. Conf. Commun.*, ICC'84, Jun., Amsterdam, The Netherlands, 1984, pp. 919-923

[WIL 1986]

S.G. Wilson, "Rate 5/6 Trellis-Code 8-PSK", *IEEE Trans. Commun.*, COM 34, n.10, Oct., 1986, pp.1 045-1049

[WIL 1986]

S.G. Wilson, "Bandwidth-Efficient Modulation and Coding: A Survey of Recent Results", *Proc. Int. Conf. Commun.*, ICC'86, Jun., Toronto, Canada, 1986, pp. 31.1.1-31.1.5

[WOL 1978]

J.K. Wolf, "Efficient Maximum Likelihood Decoding of Linear Block Codes Using a Trellis", *IEEE Trans. Info. Theory*, IT 24, n.1, Jan., 1978, pp. 76-80

[WOL&UNG 1986]

J.K. Wolf and G. Ungerboeck, "Trellis Coding for Partial-response Channels", *IEEE Trans. Commun.*, COM 34, n.8, July, 1986, pp. 751-772

[WON&McL 1988]

L.-N. Wong and P.J. McLane, "Performance of Trellis Codes for a Class of Equalized ISI Channels", *IEEE Trans. Commun.*, 36, n.12, Dec, 1988, pp. 1330-1336

[WOZ&JAC 1967]

J.M. Wozencraft and I.M. Jacobs, "Principles of Communication Engineering", New York: Wiley, 1965

[WU et al. 1987]

W.W. Wu, D. Haccoun, R. Peile and Y. Hirata, "Coding for Satellite Communication", *IEEE J. Select. Areas Commun.*, SAC 5, n.4, May, 1987, pp. 724-747

[WYN 1965]

A.D. Wyner, "The Capacity of the Band-Limited Gaussian Channel", *Bell Syst. Tech. J.*, 45, Mar., 1965, pp. 359-371 in: Key Papers in "The Development of Info. Theory", D. Slepian (Ed.), New York: IEEE Press, 1973

Y

[YUE et al. 1990]

J.H. Yuen, M.K. Simon, W. Miller, F. Pollara, C.R. Ryan, D. Divsalar and J. Morakis, "Modulation and Coding for Satellite and Space Communications", *Proc. IEEE*, 78, n.7, July, 1990, pp. 1250-1266

Z

[ZEH&WOL 1987]

E. Zehavi and J.K. Wolf, "On the performance Evaluation of Trellis Codes", *IEEE Trans. Info. Theory*, IT 33, n.2, Mar., 1987, pp. 196-202

[ZEH&WOL 1988]

E. Zehavi and J.K. Wolf, "On Saving Decoder States for Some Trellis Codes on Partial Response Channels", *IEEE Trans. Commun.*, COM 36, n.2, Feb., 1988, pp. 222-224

[ZHA&VUC 1989]

L. Zhang and B. Vucetic, "Block Coded Modulation on AWGN and Fading Channels", presented at *IEEE Int. Symp. Info. Theory, ISIT'90*, San Diego, USA, Jan. 1990, p. 126

[ZHA 1992]

W. Zhang, (en préparation) "*Combinaison de Codes de Reed-Solomon avec la Modulation MDP*", Thèse de Doctorat de l'ENST, TELECOM Paris, 1992

[ZHA&BAT 1992]

W. Zhang and G. Battail, "On the Euclidean Weight Distribution of Coded-Modulation Combining MDS codes over $GF(q)$ and q -ary Modulation", non publié

[ZWE 1990]

S. Zweig, "Le premier mot qui traversa l'océan", in: *Les riches heures de l'humanité*, éditions Belfond, Paris, 1989

RELATION DES TRAVAUX PUBLIES

Le contenu de cette thèse a fait l'objet des articles et communications suivants:

- [1] G. Battail, H.M. de Oliveira et W. Zhang, "Coding and Modulation for the Gaussian Channel in the Absence or in the Presence of Fluctuations", *EUROCODE'90 proceedings*, G. Cohen and P. Charpin (Eds.) Berlin: Springer-Verlag, 1991, pp. 337-349
- [2] G. Battail et H.M. de Oliveira, "Probabilité d'erreur du codage aléatoire avec décodage optimal sur le canal additif gaussien, affecté ou non de fluctuations d'amplitude", submitted to *Ann. Telecomm.*
- [3] H.M. de Oliveira et G. Battail, "A Capacity Theorem for Lattice Codes on Gaussian Channels", *Proc. IEEE/ST Int. Telecomm. Symp. ITS'90*, Rio de Janeiro, Brazil, 1990, pp. 1.2.1-1.2.5
- [4] H.M. de Oliveira et G. Battail, "The Random Coded-Modulation: Performance and Euclidean Distance Spectrum evaluation", *Proc. IEEE Int. Symp. Info. Theory ISIT'91*, Budapest, Hungary, June 1991, p. 2
- [5] H.M. de Oliveira et G. Battail, "The Random-Coded Modulation: Performance and Distance Spectrum Evaluation", à paraître, *Ann. Telecomm.*
- [6] H.M. de Oliveira et G. Battail, "On Generalized Constellations and the Opportunistic Secondary Channel", proposé aux *Ann. Telecomm.*
- [7] H.M. de Oliveira, W. Zhang et G. Battail, "Performance of Lattice Codes on the Gaussian Channel", submitted to *ISSSE'92*

Sur la Théorie des Systèmes de Modulation Codée pour la Transmission Numérique à Débit Elevé.

Le début est consacré à un aperçu historique et le reste à des contributions personnelles. Le problème de reconstituer la phase de la porteuse a été envisagé pour les Modulations Codées en Treillis (MCT), aboutissant à une MCT différentiellement cohérente. Un étiquetage binaire optimal des symboles d'une constellation carrée a été introduit et une famille de constellations en croix généralisées étudiée. Des bornes fines de leurs performances ont été établies et leur efficacité spectrale a été évaluée.

Utilisant le théorème de géométrie des nombres de Minkowski-Hlawka, on a démontré un théorème de capacité pour le canal à bruit additif, blanc et gaussien, établissant ainsi l'existence de moyens déterministes pour approcher la capacité.

Des bornes sur les performances des codes de réseau ont été obtenues. Une amélioration de l'approximation continue de Forney pour estimer la puissance moyenne normalisée à 2 dimensions a été suggérée. On a analysé le comportement de ces bornes pour un grand nombre de dimensions et conjecturé que s'approcher de la capacité est lié à la recherche de bons recouvrements plutôt qu'à celle d'empilements denses.

On a appliqué le raisonnement de Shannon basé sur le codage aléatoire à des constellations de nombre de dimensions fini. Les spectres de distance euclidienne de telles constellations ont été calculés. Un phénomène de durcissement de la distance a été trouvé. On a montré que l'affirmation "tous les codes sont bons" reste vraie pour les constellations codées.

ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS

CENTRE DOCUMENTAIRE - IMPRIMERIE - 46 rue Barrault - 75634 Paris Cedex 13

Téléphone (1) 45 81 77 77 - Télécopie (1) 45 89 79 06 - Telex SUPTLCM 200160F

©1992 FRANCE TELECOM - RCS PARIS B 380 129 866



DIRECTION DE L'ENSEIGNEMENT SUPERIEUR